DOCUMENT RESUME

ED 378 145 SP 035 663

AUTHOR	Devaney, Barbara; And Others
TITLE	Evaluating the Effects of School Health Interventions on School Performance. Design Report.
INSTITUTION	Mathematica Policy Research, Princeton, N.J.
SPONS AGENCY	Public Health Service (DHHS), Rockville, MD. Office of Disease Prevention and Health Promotion.
REPORT NO	MPR-8086-100
PUB DATE	1 Aug 93
CONTRACT	282-92-0044
NOTE	114p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Academic Achievement; Elementary Secondary Education; Evaluation Research; Food Service; *Health Education; *Health Programs; Health Promotion; Health Related Fitness; *Intervention; Outcomes of Education; Physical Education; Program Effectiveness; *Public Health; Research Design; *School Activities; School Health Services; Schools; Student Attitudes

ABSTRACT

This evaluation design report presents a general framework for assessing the effects of school health interventions on students' school performance in order to guide efforts to develop strong empirical evidence. The report begins with an overview of eight general types of school health interventions: school health education, health services, efforts to promote a healthy school environment, school food services, physical education and fitness, integration of school and community health promotion, school counseling, and health promotion for school faculty and staff. The report then reviews measures of school performance and primary data sources for obtaining those measures. Measures include education outcomes, student behaviors, and student attitudes. The primary data sources are school records and student surveys. The report reviews the procedures needed to collect data from each source. The report concludes by reviewing three major evaluation design options: (1) encourage and support the evaluation efforts of school districts; (2) exploit the data collected by the national surveys directed by the National Center for Education Statistics; and (3) conduct multi-school demonstrations either as add-ons to existing evaluations or as new efforts to test alternative health interventions. Contains 81 references. (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

SP

Contract No.: 282-92-0044
MPR Reference No.: 8086-100

ED 378 145

EVALUATING THE EFFECTS OF SCHOOL HEALTH INTERVENTIONS ON SCHOOL PERFORMANCE

DESIGN REPORT

August 1, 1993

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Authors:

Barbara Devaney
Peter Schochet
Craig Thornton
Nancy Fasciano
Amelia Gavin

Submitted to:

U.S. Department of Health and Human Services
Office of Disease Prevention and Health Promotion
Switzer Building, Room 2132
330 C Street, SW S.W.
Washington, DC 20201

Project Officers:
James Harrell
Robert St. Peter

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
(609) 799-3535

Project Director:
Barbara Devaney

SP035663

BEST COPY AVAILABLE

ACKNOWLEDGMENTS

In developing this design for studying the effects of school health interventions on school performance, we benefited from the advice and assistance of J. Michael McGinnis, Director of the Office of Disease Prevention and Health Promotion (ODPHP), James Harrell and Robert St. Peter of ODPHP, and from an advisory panel that included:

- Robert Boruch, University of Pennsylvania
- Claire Brindis, University of California, San Francisco
- Joseph Caliguro, U.S. Department of Education
- Janet Collins, Centers for Disease Control and Prevention
- Richard Duran, University of California, Santa Barbara
- William Hansen, Wake Forest University, North Carolina
- Richard Jung, The Bullis School, Washington, DC
- Joel D. Killen, Stanford University
- Richard Murnane, Harvard University
- Bruce Simons-Morton, National Institute of Child Health and Human Development
- Elaine Stone, National Heart, Lung, and Blood Institute
- Gene Unger, Santa Clara Unified School District

We also benefited from the advice of several of our colleagues at Mathematica Policy Research, Inc. who have been evaluating programs and strategies to improve school performance. In particular, Ellen Kisker, David Myers, Mary Moore, Mark Dynarski, Alan Hershey, and John Love provided valuable assistance. Finally, the report benefited greatly from the efforts of Laura Berensen, who edited the report, and Cindy Castro, who directed its production.

CONTENTS

Chapter		Page
	ACKNOWLEDGMENTS	iii
	ABSTRACT	ix
I.	INTRODUCTION	1
	A. RATIONALE FOR THE STUDY	2
	B. DESIGNING AN EVALUATION OF THE EDUCATION OUTCOMES OF SCHOOL HEALTH INTERVENTIONS	4
	C. OVERVIEW OF THE DESIGN REPORT	5
II.	SCHOOL HEALTH INTERVENTIONS	7
	A. CONCEPTUAL MODEL	8
	B. SCHOOL HEALTH INTERVENTIONS	12
	1. School Health Education	12
	2. School Health Services	16
	3. Healthy School Environment	20
	4. School Food Service	21
	5. Physical Education and Fitness	22
	6. Integrated School and Community Health Promotion	23
	7. School Counseling	24
	8. Health Promotion for School Faculty and Staff	25
III.	SCHOOL PERFORMANCE MEASURES	27
	A. MEASURES OF EDUCATION OUTCOMES	27
	B. MEASURES OF EDUCATION BEHAVIORS	32
	C. STUDENT ATTITUDES	34
	D. DESCRIPTIVE INFORMATION	37
IV.	OBTAINING SCHOOL PERFORMANCE MEASURES	41
	A. SCHOOL RECORDS DATA	41
	1. Obtaining School District Cooperation	42
	2. Obtaining Informed Consent	43
	3. Collecting Student Records Data	45
	4. Obtaining Supplemental Data Items	47
	B. NATIONAL DATA SETS	54

CONTENTS *(continued)*

Chapter		Page
V.	EVALUATION DESIGN OPTIONS	63
	A. EVALUATIONS OF LOCAL SCHOOL HEALTH PROGRAMS	64
	B. ANALYSIS OF EXISTING DATA	68
	1. Analysis of Secondary Survey Data	68
	2. Analysis of Data from Previous Evaluations of School Health Interventions	78
	C. LARGE-SCALE, MULTISITE DEMONSTRATIONS	81
	1. The Case for Random Assignment	82
	2. Sample Design for New School Health Interventions	83
	3. Sample Design for Existing School Health Interventions	92
	4. Process Analysis	95
	REFERENCES	97

TABLES

Table		Page
II.1	OVERVIEW OF SELECTED SCHOOL HEALTH EDUCATION PROGRAMS	13
III.1	DATA REQUIREMENTS FOR MEASURING STUDENT SHORT-TERM EDUCATION PERFORMANCE	28
III.2	REPRESENTATIVE ACHIEVEMENT TESTS	31
IV.1	STUDENT SECONDARY SCHOOL EDUCATION PERFORMANCE MEASURES IN THE NLS72, HS&B, AND NELS DATABASES	56
V.1	PRELIMINARY SPECIFICATIONS OF VARIABLES TO BE COLLECTED BY THE EARLY CHILDHOOD LONGITUDINAL SURVEY	71
V.2	MINIMUM DETECTABLE IMPACTS ON VARIOUS EDUCATION PERFORMANCE MEASURES, USING A RANDOM-SAMPLE DESIGN	89
V.3	MINIMUM DETECTABLE IMPACTS ON VARIOUS EDUCATION PERFORMANCE MEASURES, USING A TWO-STAGE, CLUSTERED DESIGN	90

FIGURES

Figure		Page
II.1	A SIMPLE MODEL OF THE LINKS BETWEEN SCHOOL HEALTH PROGRAMS AND STUDENT SCHOOL PERFORMANCE	9
IV.1	DROPOUT DEMONSTRATION ASSISTANCE PROGRAM: STUDENT RECORDS FORM	48

ABSTRACT

EVALUATING THE EFFECTS OF SCHOOL HEALTH INTERVENTIONS ON SCHOOL PERFORMANCE: DESIGN REPORT

August 1, 1993

Barbara Devaney, Peter Schochet, Craig Thornton, Nancy Fasciano, Amelia Gavin
Mathematica Policy Research, Inc.
Princeton, New Jersey

This evaluation design report presents a general framework for assessing the effects of school health interventions on school performance. Health and education appear to be inextricably linked: good health is necessary for effective learning, and education is necessary for maintaining good health. However, robust empirical evidence from well-designed evaluations of the link between health and education outcomes for children and youth currently does not exist. This study was undertaken in order to guide efforts to develop such strong empirical evidence.

The report begins with an overview of eight general types of school health intervention: (1) school health education; (2) health services; (3) efforts to promote a healthy school environment; (4) school food services; (5) physical education and fitness; (6) integrated school and community health promotion; (7) school counseling; and (8) health promotion for school faculty and staff.

The report then reviews measures of school performance and primary data sources for obtaining those measures. Measures include education outcomes (e.g., graduation rates, grade promotion patterns, grades, and standardized achievement tests), student behaviors (e.g., attendance and drop-out status), and student attitudes (e.g., self-esteem and locus of control). The primary data sources are school records and student surveys. The report reviews the procedures needed to collect data from each source.

The report concludes by reviewing three major evaluation design options. The first option is to encourage and support the evaluation efforts of school districts. The second is to exploit the data collected by the national surveys directed by the National Center for Education Statistics. The third is to conduct multischool demonstrations, either an add-ons to existing evaluations of the health outcomes of school health programs, or as new efforts to test alternative health interventions.

I. INTRODUCTION

Fueled by concerns about substance abuse, sexually transmitted diseases, and a vast array of social, emotional, and physical health problems with which students must contend, school health programs have become an important part of school curricula during the past decade. Although the objectives of school health programs vary greatly, they include: (1) developing and implementing a comprehensive, sequential school health curriculum for grades kindergarten through 12; (2) promoting health maintenance and wellness, rather than simply preventing disease; (3) developing students' decision-making skills and sense of individual responsibility for their own health; (4) improving students' health-related knowledge, attitudes, and practices; and (5) integrating the physical, mental, emotional, and social dimensions of health (Allensworth 1993).

Despite evidence that the health needs of students are both increasing and becoming more complex, school health programs are not yet firmly rooted in public school systems. The nation's public schools are under enormous pressure to improve the academic skills of students, especially in the core subject areas of language and mathematics. This pressure has resulted in the fairly widespread attitude among school officials that school health programs are less critical to academic success and may, in some cases, consume time that could otherwise be devoted to academic instruction. In addition, as school administrators respond to severe budgetary constraints, school health programs often are canceled or assigned to part-time or less well-trained teachers.

As school and health administrators debate the continuing development of school health programs, they will need information not only about how these programs affect health, but also information about how they affect students' school performance. Information about impacts on health comes from the numerous demonstrations and evaluations that have yielded a wealth of information about impacts on health knowledge, attitudes, behavior, and outcomes. However, analyses of the education outcomes of school-based health interventions are much fewer in number

and generally comprise only a small part of larger evaluations. Consequently, robust empirical evidence from well-designed evaluations of the link between health and education outcomes for children and youth does not exist.

To address the lack of this critical evidence, Mathematica Policy Research, Inc. (MPR) is conducting a study to Design an Evaluation of the Effects of School Health Interventions on School Performance. The study includes a literature review (Fasciano and Devaney 1993) and a review of key issues in evaluation design. The following sections discuss more thoroughly the rationale for the study and the objectives of key components.

A. RATIONALE FOR THE STUDY

Health and education are inextricably linked. Good health is necessary for effective learning; education is necessary for maintaining good health (Novello, et al. 1992). Students' readiness to learn throughout childhood depends largely on good nutrition, prevention of illnesses and injuries, protection from violence and other health hazards, and early treatment of disabilities and developmental problems. To maintain good health throughout their lives, children and youth must acquire knowledge about health risks and must develop the skills necessary to avoid or reduce those risks. Finally, adults cannot realize the long-term benefits of education--increased productivity and active participation in society--unless they are also able to maintain their health.

Our national education agenda--America 2000--and our national health agenda--Healthy People 2000--explicitly recognize the relationship between education and health. Many of the national education goals include health objectives. For example, the first education goal--that, "by the year 2000, all children in America will start school ready to learn"--will require health promotion among young children (National Education Goals Panel 1991). The third objective of this goal states that "children will receive the nutrition and health care needed to arrive at school with healthy minds and bodies." The sixth national education goal--"by the year 2000, every school will be free of drugs and

violence and will offer a disciplined environment conducive to learning"--also recognizes the relationship among effective education, good health, and safety.

The Healthy People 2000 goals include education objectives. For example, the National Health Promotion and Disease Prevention Objectives set forth in Healthy People 2000 include objectives associated with developmental preschool, health education, school absenteeism, and high school graduation rates (Novello et al. 1992). Health education is an important focus of the objectives: Objective 8.4 proposes to "increase to at least 75 percent the proportion of the nation's elementary and secondary schools that provide planned and sequential kindergarten through twelfth grade quality school health education." Recognizing that children should remain in school, one health objective calls for increasing "the high school graduation rate to at least 90 percent, thereby reducing risks for multiple problem behaviors and poor mental and physical health."

Because the national health and education goals are so interrelated, schools provide a natural setting for coordinating efforts to achieve them. According to McGinnis and DeGraw (1991), more than one-third of the Healthy People 2000 objectives can be influenced significantly or achieved directly by schools. Moreover, because most children and youths attend public school, the schools are potentially the most systematic and efficient providers of health education and health care services for children.

In fact, schools have adopted a number of strategies for promoting good health and preventing health problems. These strategies include providing health education, offering school-linked or school-based health services, creating a healthy and safe school environment, requiring or encouraging physical education, offering healthful meal choices from school food services, providing psychological assessments and counseling to promote child development and emotional health, promoting health awareness among faculty and staff on site at schools, and integrating school and community health-promotion efforts (McGinnis and DeGraw 1991). The objectives of most school health programs are to bring about immediate improvements in health-related knowledge, attitudes, and behavior,

although some programs seek to improve students' current physical or mental health, and others aim to teach skills that will enable students to recognize and resist peer pressures that might encourage unhealthy behaviors.

Achieving these interrelated health and education goals and refining these emerging school health programs, however, are limited by the categorical approach of many school health programs and the lack of evidence to document the relationship between health and education. Future evaluation efforts can help to overcome this limitation by focusing on school performance impacts, as well as on health outcomes.

B. DESIGNING AN EVALUATION OF THE EDUCATION OUTCOMES OF SCHOOL HEALTH INTERVENTIONS

Although the belief that school health programs will influence education outcomes for children and youths is widespread, little evidence is available from rigorous evaluations to support this belief. This study will begin to address the dearth of evidence by providing a framework for assessing the education outcomes associated with school health interventions. The study has three components:

- A review of the literature and evaluation issues
- Consultation with an advisory panel
- Development of an evaluation design report

The literature review had two components: (1) a summary of current knowledge about the effectiveness of school health interventions (Fasciano and Devaney 1993); and (2) an overview of design issues that are key to developing a comprehensive framework for evaluating the education outcomes of school health programs.

After completion of the literature review and identification of key evaluation design issues, an advisory panel of researchers and policymakers was convened to discuss the critical issues. The panel, with diverse interests and expertise in the areas of school health, education outcomes, and evaluation

design and analytic methods, advised the MPR study team on the types of health interventions likely to affect school performance, the education outcomes that should be considered, the availability of school records and protocols for assessing the outcomes, and the types of evaluation designs that are feasible and appropriate for school settings.

Building on the literature review and the discussion of key evaluation design issues with the advisory panel, this evaluation design report provides a general framework for assessing the education outcomes of school health interventions. It discusses the types of school health interventions that could affect school performance, relevant education outcomes to measure, alternative evaluation strategies, sample design considerations, and data collection strategies.

C. OVERVIEW OF THE DESIGN REPORT

The following chapters present a framework for evaluating the effects of school health interventions on education outcomes. Chapter II provides both a conceptual model of the relationship between health and education and an overview of the key school health interventions. Chapter III discusses concepts and measures used to assess school performance, including student attitudes, behaviors, and education outcomes. Chapter IV discusses possible data sources for an evaluation, particularly school records and surveys. Chapter V describes three possible design options: (1) an evaluation of local school health programs; (2) analysis of student survey data; and (3) large-scale, multisite demonstration evaluations.

II. SCHOOL HEALTH INTERVENTIONS

Concern about both the health status and educational achievement level of America's children has heightened during the past decade. Children today face a myriad of serious health problems. Many of the risk factors for chronic diseases during adulthood have their behavioral roots in childhood and adolescence. For example, early use of tobacco, alcohol, and marijuana has been shown to correlate with alcoholism and drug abuse during adolescence and adulthood (Clayton 1981). Lifetime patterns of diet and physical activity may be established in childhood, as well (U.S. Department of Health and Human Services, Public Health Service 1990). Early sexual activity can lead to risks of unwanted pregnancies and infection by HIV/AIDS or other sexually transmitted diseases. Furthermore, for many children, especially those from low-income households, both access to and the resources to pay for primary preventive health care are limited. Consequently, these children attend school with untreated health problems.

During the last decade, America's schools have been criticized intensely for their failure to maintain high educational standards. Scores on standardized tests have declined, alarmingly high drop-out rates persist, substance abuse and violence plague schools, and American students consistently perform poorly relative to students from other nations on standardized tests administered internationally (Chubb and Moe 1990).

Responses to the health and education problems that affect America's children typically have reflected a categorical, rescue-mentality approach. To address the health problems of America's children, a vast array of school health programs has been developed and implemented in public schools across the country. To respond to the crisis in education, a broad set of education reforms has been initiated, ranging from raising academic requirements to establishing magnet schools or implementing parental school choice. For the most part, these efforts have been evaluated in the same categorical manner: school health programs have been assessed with respect to their effects

on health-related outcomes, whereas education reforms have been evaluated for their impacts on academic outcomes.

More recently, a growing awareness of the link between the health and education problems of America's children has emerged. Children must be physically and emotionally healthy in order to learn effectively, and education is a prerequisite for learning about and leading a healthy lifestyle (Novello et al. 1992). This chapter presents an intuitive and simplified conceptual model of the relationship between student health status and educational achievement. It also briefly describes basic types of school health interventions that can affect school performance.

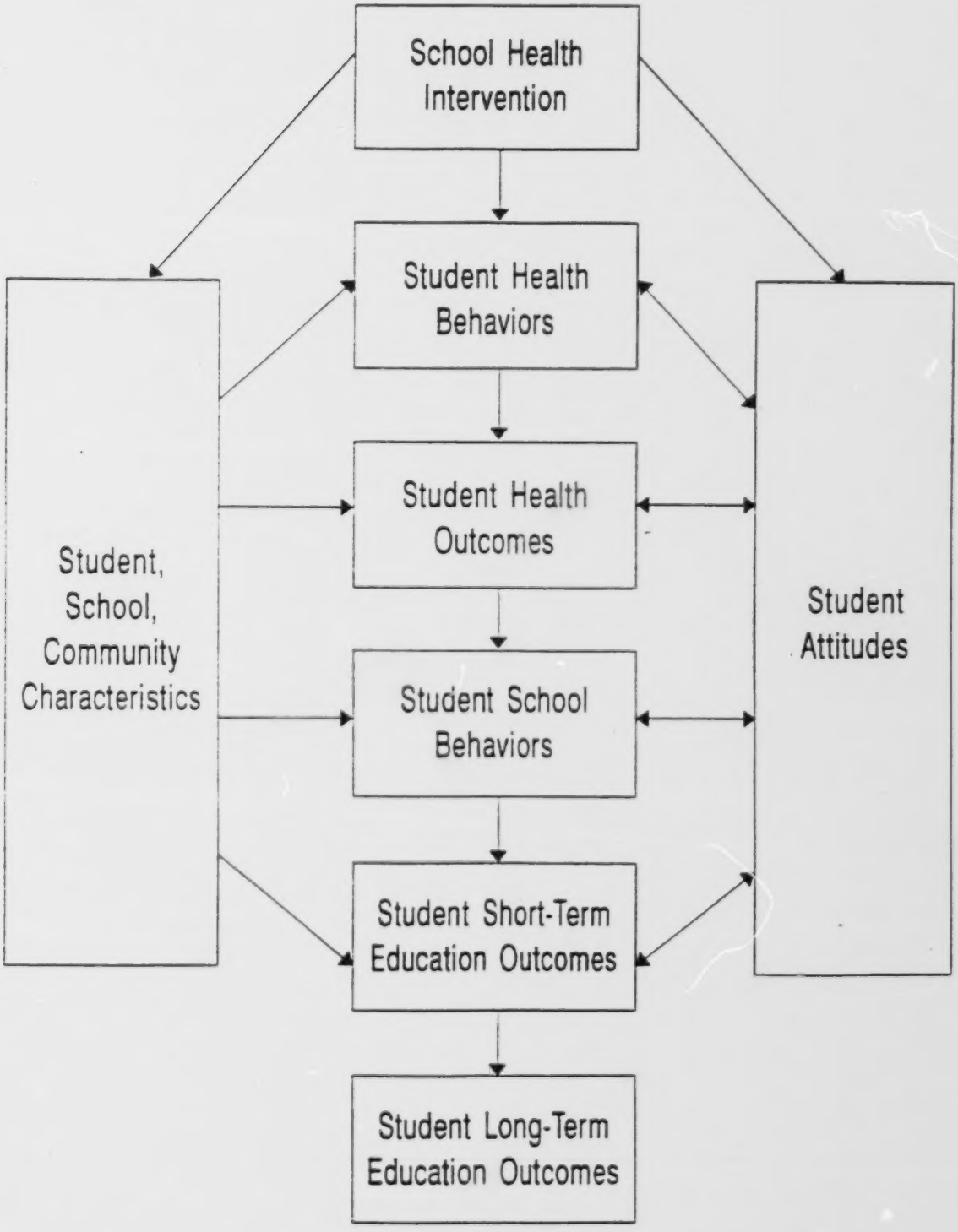
A. CONCEPTUAL MODEL

For the most part, school health programs affect student school performance through an indirect, multistage process. A simple model of this process would begin with the school health intervention directly affecting student health behaviors and attitudes, and the nature of the school environment (Figure II.1). These direct effects would then affect student health status, which, in turn, would affect student school-related behaviors. The altered student behaviors would then affect student education outcomes over the short and the long term. Throughout this process, the characteristics of the students, their families, student peer groups, and the school/community environment also work to shape the pattern and magnitude of any impacts.

A key feature of this model is the distance between the health intervention and the education outcomes. This distance may make it difficult for many of the interventions to affect the outcomes, and for evaluations to detect such effects, when they occur. This difficulty in affecting outcomes is most severe when examining school health interventions that are limited in either scope or duration. In these cases, which appear to constitute the majority of school health interventions, the effects are likely to be small and will be lost among the effects of all of the other influences on school performance. Even for programs with substantial impacts, evaluations may need a large sample and rigorous designs in order to detect impacts on school performance.

FIGURE II.1

A SIMPLE MODEL OF THE LINKS BETWEEN SCHOOL HEALTH PROGRAMS AND STUDENT SCHOOL PERFORMANCE



The model also suggests that impacts on student attitudes or behaviors can be used as proxies for longer-term impacts on education outcomes. The advantage of such proxy outcomes is that impacts may be easier to detect, given the more direct links between these intermediate outcomes and the health intervention. For example, school health interventions may have a more discernable impact on the student characteristics that they affect directly, such as health attitudes, school attendance, or disruptive actions, than on education outcomes, which they affect only indirectly. In addition, impacts on behaviors and attitudes may become evident earlier than would impacts on outcomes; thus, the former would be more appropriate targets for short-term evaluations.

The problem with using proxies stems from the fact that impacts on the intermediate proxy factors do not always imply that the education outcomes of interest are affected as well. For example, studies of substance abuse prevention programs have found that impacts on student knowledge and reported attitudes do not always translate into changes in behavior (Fasciano and Devaney, 1993). Because of the uncertainty surrounding the links between the various factors shown in Figure II.1, any final decision about using proxy indicators to assess potential impacts on education outcomes will depend on the nature of the intervention, the types of expected effects, and the available research linking the proxy outcomes to the outcomes of interest.

All of the elements in the simple model represent a diverse array of specific attitudes, behaviors, and outcomes. However, the variation in ***school health interventions*** is particularly noteworthy. In the United States, these interventions include the following eight broad categories of programs (Anderson and Creswell 1976; Stone 1990; and Kolbe 1986a and 1986b):

- School health education
- School health services
- Healthy school environment
- School food services
- Physical education and fitness

- Integrated school and community health promotion
- School counseling
- Health promotion for school faculty and staff

As discussed in Section B, some of these programs may be narrowly focused, as in the case of a health education curriculum that covers a well-defined set of topics for a specific set of students, or a school lunch program specifically targeted at decreasing the fat intake of students. In other cases, the programs may be quite broad in scope, with the services generally available in the school or functioning as part of a larger program. For example, a school nurse typically is available to all students in a school, but may provide direct services to relatively few. Similarly, efforts to improve the school environment primarily may improve the functioning of existing programs, rather than affect student performance directly. Programs also vary in duration and intensity, ranging from a simple hearing test to extensive curricula that span several grade levels.

We will use the term "intervention" to refer to the entire range of programs, even though many school health programs might more accurately be thought of as program components, rather than as a specific, detailed strategy.. In this sense, intervention will mean the specific plan, program, program component, or strategy that is being evaluated, regardless of its intensity or magnitude.

The model implicitly defines the health intervention in terms of a specific alternative. Typically, the alternative to a new health program would be the school activities that would have taken place in the absence of that program, that is, the status quo. Thus, a program to introduce a comprehensive school health education program would be evaluated in relation to the classroom instruction that the school would have provided in the absence of the new program. Interventions could also be evaluated relative to a specific alternative program, such as a comparison of two health education programs. In either case, the model assumes that changes in attitudes, behaviors, and outcomes stem from a *change* in the type of programs provided.

B. SCHOOL HEALTH INTERVENTIONS

As a foundation for the evaluation design, we first review the eight general types of school health interventions: (1) school health education; (2) school health services; (3) healthy school environment; (4) school food service; (5) physical education and fitness; (6) integrated school and community health promotion; (7) school counseling, and (8) health promotions for school faculty and staff. For each of these types of interventions, we describe the kinds of programs that have been implemented and the findings with respect to impacts on health and school performance.

1. School Health Education

Health education, which has a long history in schools, is the formal curriculum taught to students. Although the content of individual programs varies, health instruction comprises 10 basic areas: (1) community health; (2) consumer health; (3) environmental health; (4) family life; (5) growth and development; (6) nutritional health; (7) personal health; (8) prevention and control of disease; (9) safety and accident prevention; and (10) substance use and abuse (Lohrmann et al. 1987).

Table II.1 provides an overview of the four primary types of school health education programs and describes selected program curricula in each category. The first type--***comprehensive school health education programs***--typically cover several or all of the eight content areas. The other three types--***cardiovascular health education, nutrition education, and substance abuse prevention programs***--focus on only one or two areas.

Comprehensive school health programs are the most common school health interventions. They are based on the premise that children are less likely to engage in substance abuse and other unhealthy behaviors, such as consuming high-fat diets or being physically inactive, if they understand how their bodies work, develop a strong sense of self and good decision-making skills, and take personal responsibility for their health. The rationale for both nutrition education programs and cardiovascular fitness programs is that unnecessary morbidity, premature mortality, and chronic adult diseases are caused by specific behaviors, which are often established in childhood and adolescence.

TABLE II.1

OVERVIEW OF SELECTED SCHOOL HEALTH EDUCATION PROGRAMS

Comprehensive School Health Education Programs	
School Health Curriculum Project (SHCP)	Targets a different body system (digestive, respiratory, cardiovascular, and central nervous) at each grade level (4 to 7). Each curricular unit is covered one class period per day, five days per week, over a 10- to 12-week period.
Health Education Curriculum Guide (HECG)	A comprehensive health program implemented in grades K-6 focusing on six areas of health education: (1) nutrition; (2) family living; (3) drugs, alcohol, and tobacco; (4) human growth and development; (5) environmental, community, and mental health; and (6) personal safety. HECG takes a "cookbook" approach, permitting teachers and school districts to select and teach activities/lesson plans in each of the six areas in any order.
Project Prevention	An integrated health curriculum for grades K-9 and 11, focusing on such areas as growth and development, self-concept, personal health habits, basic first aid, decision-making skills, and community health. The curriculum consists of activities that can be conducted throughout the school year. When these activities are fully implemented, they consume about 8 percent of total instructional time.
Teenage Health Teaching Modules (THTM)	Designed to be compatible with the SHCP elementary school health curriculum, this program consists of 16 modules and seeks to develop five skills: (1) self-assessment; (2) communication; (3) decision making; (4) health advocacy; and (5) healthy self-management.
Cardiovascular Health Education Programs	
Reading', Riting', Rithmetic, and High Blood Pressure	A limited-focus and limited-duration program that attempts to educate students about the causes and dangers of high blood pressure. It is a 10-session program for sixth graders, combining instruction on the circulatory system with training in using a stethoscope and blood-pressure cuff, and including a home practicum to promote discussion among parents and children.
Know Your Body	A multigrade curriculum that focuses on the prevention of heart disease and includes medical screening for such risk factors as elevated cholesterol levels, obesity, cigarette smoking, and high blood pressure. The curriculum teaches both basic health skills and individual responsibility for health.
Child and Adolescent Trial for Cardiovascular Health	A program for grades 3-5 that includes classroom curricula and school environmental modifications related to food consumption, physical activity, and tobacco use in order to influence favorably blood concentrations of lipids, blood pressure, body fat, and physical fitness.

TABLE II.1 *(continued)*

Nutrition Education Programs	
Nutrition in a Changing World	A sequential curriculum for grades K-12 that promotes general nutrition knowledge and encourages an appreciation of a variety of foods. The curriculum focuses on the concept of food as a source of nutrients and on the function of nutrients relative to health.
Hearty Heart and Friends	A 10-week, 20-session curriculum for elementary school children. Based on social learning theory, which holds that behavior is the product of the interaction among personal, environmental, and behavioral characteristics, the curriculum is narrowly focused, targeting specific behavior changes. Designed to change children's fat, salt, and complex carbohydrate consumption, the program includes one session per week in which new information is introduced, and another weekly session in which students prepare nutritious snacks.
Great Sensations	A six-session program for high school students that draws on social learning theory and emphasizes modeling, behavioral rehearsal, and reinforcement. In promoting healthy, low-salt snacking, class sessions include a range of participatory learning activities. Students study the health consequence of high-salt consumption and learn to identify salty snack foods, decipher product labels, and reduce their intake of salty snacks. Along with the classroom component, Great Sensations includes a school media campaign to reinforce health messages. Parents may also be contacted to urge them to support their children's use of low-salt snack foods.
Substance Abuse Prevention Programs	
Social Influence Programs: Project CLASP High School Smoking Prevention Program Waterloo Smoking Prevention Program Alcohol Misuse Prevention Study Adolescent Alcohol Prevention Trial	Programs designed to prevent substance abuse by: (1) educating students about social influences; (2) strengthening their skills for resisting those influences; and (3) correcting or establishing perceptions of social norms.
Personal and Social Skills Model Social Skills Training Program Life Skills Training Program	Programs that help students acquire personal and social skills to use in situations involving substance abuse. Programs typically stress problem-solving and decision-making skills, assertiveness, stress-reduction strategies, and conversational and other social skills.

These programs typically seek to reduce the fat, sodium, and cholesterol content of diets and to prevent the excessive intake of calories. The cardiovascular fitness programs may also seek to increase physical activity and to delay or prevent the onset of smoking or the transition from experimental to chronic tobacco use. Finally, substance abuse prevention programs, which are based on theoretical models of adolescent development and behavior, attempt to prevent substance abuse through education and development of skills to resist peer pressure.

Comprehensive health education programs are implemented in secondary schools, as well as elementary schools, whereas most nutrition and cardiovascular education programs are targeted to elementary school students. Conversely, many substance abuse prevention programs are aimed at junior and senior high school students.

All of the programs listed in Table II.1 have been evaluated with respect to their health outcomes. In most instances, program effectiveness was measured in terms of improvements in students' health knowledge, attitudes, and self-reported behaviors. Although results vary somewhat across studies, school health education programs typically are shown to produce moderate to large gains in students' health knowledge, and smaller gains in improvements in attitudes and behaviors. Evaluations of nutrition and cardiovascular health programs that actually observed student behavior (such as food consumption in the cafeteria), or that assessed risk factors for heart disease, indicate that the programs had varying effectiveness. The Know Your Body program, for example, was associated with favorable changes in some disease risk factors, including high blood pressure and cholesterol levels, but had no impact on obesity. Program effectiveness may also vary with the age of the students or over time. For example, some nutrition programs were found to be less successful at influencing the behavior of children in the upper elementary grades than in the lower grades; some behavioral changes were also short-lived. The effectiveness of substance abuse prevention programs in changing behaviors are also varied. Although it is difficult to draw general conclusions about the

effectiveness of these programs, ongoing intervention appears to be necessary in order to sustain positive behavioral effects (see, for example, Ellickson, et al. 1993; and Flay et al. 1989).

The theory that improving students' health status or behaviors will improve school performance has prompted a few evaluations of school health education programs to include education outcomes among their measures of effectiveness. These measures, which include achievement test scores, grade point averages, absenteeism, teachers' reports of classroom misbehavior, and reports of school disciplinary actions, are somewhat more likely to be used to evaluate substance abuse prevention programs than other program types. However, substance abuse prevention programs directed toward all students generally have had disappointing results; the small number of programs that have been evaluated with respect to education outcomes were found to have little, if any, impact on these outcomes (see, for example, Schaps et al. 1986; or Pentz and Tolan 1985).

Programs that address a cluster of risk-taking behaviors, and that target students at high risk for substance abuse, dropping out, delinquency, or other deviant behaviors, may be more effective. One such program, a single-semester, psycho-educational counseling class for at-risk high school students, had very positive results (Eggert et al. 1990). The program helped students to remain in school, improve their grade point averages, decrease their use of drugs, and reduce the number of disciplinary actions to which they were subject.

2. School Health Services

School health services have been part of school health programs since the beginning of this century. Health services generally have three components: (1) health assessments; (2) preventive measures; and (3) corrective measures. Health-assessment practices vary widely and may include vision and hearing testing, height and weight measurements, physical health examinations, and dental examinations, whereas preventive measures include safety precautions, health counseling and instruction, and communicable disease control. Corrective services may include early interventions, emergency care and first aid, triage of illness and injuries, appropriate referrals, follow-up assessments

and services, and correction of problems (Stone 1990). The Individuals with Disabilities Act (PL 101-476) mandates that schools also provide any special health services required by children with disabilities to benefit from special education.

In recent years, ***school-based clinics*** and ***school-linked clinics*** have been gaining popularity as a vehicle for providing health care and promoting healthy lifestyles among adolescents. School-based health care has become an increasingly important addition to traditional health care sources, especially for adolescents who face financial or institutional barriers to other sources of care. Many school-based clinics are located in junior or senior high schools that have high rates of dropout, school failure, absenteeism, and adolescent pregnancy, and whose students suffer from multiple health problems. Approximately 90 percent of the school-based clinics are located within the main school building, and 10 percent are in a separate building on the school grounds. School-linked clinics are located close to school grounds and typically serve more than one school. Most clinics are open during the entire school day, five days a week. About 50 percent of the school-based clinics and 93 percent of school-linked clinics maintain summer operations.[1]

Approximately 90 percent of school-based clinics offer primary health care and physical examinations, as well as treatment for minor acute illnesses, accidents, and injuries. More than three-fourths of the clinics provide individual mental health, psychosocial, sexuality, family, and job counseling. In addition, about 90 percent perform laboratory tests, provide nutrition and health education, and refer students with other problems to counseling and community health care services.

Most students visit the clinics to receive physical examinations, or for treatment of acute illnesses or minor emergencies (Millstein 1988). Although media reports have focused on the birth control services provided by school-based clinics, only 10 percent to 20 percent of visits are for family planning services. About 15 percent of the clinics distribute birth control pills on site, and about 18

[1]This profile of school-based and school-linked clinics is based on a February 1991 mail and telephone survey conducted by the Center for Population Options (Waszak and Neidell 1991). Approximately 327 clinics were contacted. Of these, 239 (224 school-based clinics and 15 school-linked clinics) responded to a detailed questionnaire covering clinic activities and organization.

percent dispense condoms. Roughly 28 percent prescribe birth control methods; two-thirds refer students to collaborating agencies for contraceptives (Waszak and Neidell 1991).

Most clinics are staffed by a multidisciplinary team of nurse practitioners and social workers or counselors; larger programs also employ physicians. A registered nurse is usually a student's first contact with the school-based clinic (Newton 1979). For routine medical problems, students are often referred to a nurse practitioner or physician's assistant, who can assess, manage, and follow up health problems, health-related behavior, and health-risk factors.[2] Students who experience psychological, emotional, or substance abuse problems or unplanned pregnancies are often referred to social workers. Social workers administer psychosocial-risk assessments in order to determine the likelihood of adverse consequences and design prevention plans that can include individual or family counseling, mentoring, or family-planning services. Physicians play an important role by providing examinations, consultations, and collaborative case management, as appropriate (U.S. Department of Health and Human Services 1986). Physicians also function as school health care planners, consultants to school health professionals, and supervisors of school health programs (Silver 1976).

An important rationale for conducting health screening in the schools is the belief that early identification of problems can help to remove many potential barriers to learning. However, research suggests that screening is not entirely effective in ensuring that students obtain care for newly detected problems: first, many physician referrals are never followed up, and second, a surprising number of follow-ups fail to confirm the screening diagnosis (Brink and Nader 1984).

---

[2]Nurse practitioners can provide a wide range of health-care services. These services may include making accurate assessments of physical health status (health histories); carrying out basic developmental, cognitive, and physical examinations, including neurological examinations that incorporate screening for perceptual difficulties; tracking students' emotional, physical, and cognitive growth patterns; counseling students on family and peer relationships and collaborating with students, families, physicians, teachers, and other school personnel to provide counseling and anticipatory guidance for children with health, developmental, or educational problems; dispensing prescribed medications and supervising treatment regimens in the school; collaborating in designing preventive health care, nutrition, and screening programs; and evaluating nursing services in the school. School nurse practitioners can also perform a variety of laboratory procedures (such as throat and urine cultures and pregnancy tests) as part of health evaluations. The extent to which they use these tests varies according to the particular health needs of the students and the availability of follow-up resources within the community.

These findings suggest that the most effective approach may be one that integrates health assessment with preventive and corrective services. Research on school-based and school-linked clinics provides some support for an integrated approach. For example, some evidence suggests that school-based clinics may promote lower hospitalization rates and improved general health status among youths (Dryfoos and Klerman 1987; and Simkin 1987). On-site health care also enables students to obtain services without leaving school for the entire day.

Some evaluations provide evidence that school-based clinics reduce substance abuse among students in the schools that they serve (Kirby et al. 1989; and Dryfoos and Klerman 1987). Findings on contraceptive practices are mixed; improvements in the use of birth control have been observed at some schools, but not at others. Similarly, reductions in birth rates have been observed at some schools that provide clinic services, but not at others (Dryfoos 1985; and Kirby et al. 1989).

Evidence on education outcomes, such as absenteeism and academic achievement, is sparse. At some schools with clinics, the absentee rates of students who used the clinics were lower than those of students who did not (Simkin 1987); at other schools, the opposite was true. At one Los Angeles high school, for example, frequent use of the clinic was associated with higher absentee rates, which, in turn, were associated with lower grade point averages. However, these results may simply reflect the poorer health status of those who used the clinic relative to those who did not.

The paucity of data and the variability of findings across studies make it difficult to draw general conclusions about how the availability of clinic services affects the health or school performance of students. To date, most studies have collected only utilization data. Relatively few have investigated whether improved access to health care translates into improved health or education outcomes.

3. Healthy School Environment

The concept of a healthy school environment embraces both the physical surroundings and psychological climate of schools. Factors affecting the physical school environment include: (1) the safety of the school location and surrounding community; (2) the condition of the school building, including biological, chemical, and physical hazards; and (3) precautions, especially those taken for sporting and recreation events, to reduce accidents and injuries (Rowe 1987). In addition, smoke-free policies in schools reinforce what students learn about the health hazards associated with smoking and exposure to environmental smoke, promote a smoke-free environment as the norm, and limit the opportunity for students to smoke on school grounds.

The psychological climate of a school involves the well-being of students, faculty, and staff. Factors contributing to schools' psychological climate include constructive interpersonal relationships, school efforts to improve the quality of teacher-student relationships, and systemic analysis and problem-solving procedures (Schultz et al. 1987).

Research suggests that safety precautions can play a large role in keeping students in school. Injury rates among school-age children are very high, and although many accidents occur off the school grounds, some researchers estimate that over one-half of all injuries are connected with school life (McKenzie and Williams 1982). As a result of these and nonschool-related accidents, children between the ages of 6 and 16 lose about 14 million school days per year (Rowe 1987). Programs to improve school safety may exist; however, we are unaware of any evaluations of such programs.

Several studies have examined the effects of school policy on student smoking. The research indicates that a regularly enforced policy restricting student smoking on or near the school grounds discourages adolescent smoking (Pentz et al. 1989). Significantly more students are smokers in schools providing on-campus smoking areas than in schools that do not provide such areas (Hawkins and Catalano 1990). Another study found that permissive smoking policies for school employees are

also associated with increased student smoking rates (Olds and Eddy 1986). None of these studies explored possible links between student smoking and education outcomes, such as absenteeism.

The school health literature on administrators' efforts to alter the general climate of their schools is sparse. Many of these efforts are too diffuse to qualify as interventions, and few can be categorized clearly as health programs. Among the efforts are cooperative learning activities, social skills training, teacher training, creative approaches to discipline, and systemic analysis and problem-solving procedures (Schultz et al. 1987). Altering classroom organization may also come under this heading. For example, the recognition that the transition from elementary school to junior high school can be stressful, and may put students at risk for health-compromising behaviors, has prompted some schools to reorganize classroom structures to ease the transition. In one restructuring, for example, new students took primary academic subjects in clustered classrooms, but had a homeroom base; the reorganization was found to reduce absenteeism and school drop-out rates and to improve bonding to the school (Hawkins and Catalano 1990).

4. School Food Service

The *Dietary Guidelines for Americans* are a broad set of recommendations for healthy dietary practices. These guidelines recommend that Americans: (1) eat a variety of foods; (2) maintain desirable weight; (3) avoid too much fat, saturated fat, and cholesterol; (4) eat foods with adequate starch and fiber; (5) avoid too much sugar; (6) avoid too much sodium; and (7) drink alcoholic beverages in moderation. The guidelines specify that individuals should derive a maximum of 30 percent of food energy from fat, and a maximum of 10 percent from saturated fat.

For many children, school meals are a significant contribution to their daily nutrient intake. Although many schools have incorporated the *Dietary Guidelines* into their meal-planning practices, school meals generally remain high in fat, saturated fat, and sodium (St. Pierre et al. 1992). In addition, the nutrient content of these meals often contrasts sharply with the students' expectations resulting from education on healthy eating habits.

Programs that are designed both to educate students about good nutrition and health and to modify school food menus generally have produced positive program-related changes in students' reported eating behaviors (Fasciano and Devaney 1993). Evidence of a link between nutrition and school performance is scarce, yet some studies suggest that school nutrition programs may affect the academic performance of children (Pollitt et al. 1978). One study of the School Breakfast Program, for example, found that low-income students who participated in the program scored higher on basic skills tests and were less likely to be tardy or absent than were students who did not participate (Meyers et al. 1989). There is no clear evidence, however, that such impacts are widespread or pertain to other groups of students.

5. Physical Education and Fitness

Concern about the low level of physical activity of school children has increased during the past decade, stimulated largely by rapidly accumulating evidence of the association between sedentary habits and coronary heart disease in adults. Although a corresponding relationship with respect to children has not been established, many health researchers argue that "the heart attack of middle age begins in childhood," as children begin developing lifelong patterns of activity and diet (Kuntzleman and Drake 1984).

School curricula traditionally have included physical education (PE). Currently, about 97 percent of public school children in grades 1 through 4 are enrolled for an average of 100 minutes per week in PE programs. However, the amount of moderate to vigorous physical activity of children in these classes may be less than recommended (Simons-Morton et al. 1991). The shortcomings of traditional PE classes have spurred the development of programs designed specifically to improve children's *health-related* physical fitness. This measure (which differs from the standard definition of fitness) focuses on four components that are known to be associated specifically with disease prevention or maintenance of a good functional capacity for day-to-day living. The programs target one or more

of the following: (1) cardiorespiratory endurance; (2) body composition; (3) muscular strength and endurance; and (4) flexibility (Pate 1991).

Some physical activity programs are implemented in combination with health education and school food service modifications. For example, *Go for Health* is a program that combines fitness-oriented physical education with a health education curriculum and with school lunches that are low in fat and sodium. The exercise component, Children's Active Physical Education, consists of five six- to eight-week units designed to encourage moderate to vigorous physical activity during PE classes. Each unit includes two or three main cardiovascular fitness activities, such as dancing, running, and jumping rope (Simons-Morton et al. 1991).

Research results are encouraging. Findings from several demonstration projects conducted in elementary and junior high schools indicate that physical activity programs can increase children's activity level and improve their cardiovascular health (Simons-Morton et al. 1991; Gilliam et al. 1982; and Kuntzleman and Drake 1984). Programs geared toward improving strength and flexibility also appear to be effective with students at the elementary and junior high school levels (Duncan et al. 1983). At least one study has found that physical activity programs can boost students' self-esteem (Kutzleman and Drake 1984). (Although the program had no apparent impact on girls, the self-esteem scores of boys enrolled in the program increased significantly.) Whether the physiological and psychological gains associated with these programs translate into better attendance or higher academic achievement is not clear, as none of the evaluations assessed education outcomes.

6. Integrated School and Community Health Promotion

Collaboration between the school and community increasingly is being recognized for its potential to promote student health. Typically, collaborative programs involve establishing either a school health council/advisory board or an interagency coalition of school and community agencies that are concerned with student health (Killip et al. 1987).

The underlying premise of integrated school and community health programs is that these initiatives offer the most effective means of providing an "ideal" school health program. This type of program: (1) provides an integrated, sequential K-12 health education program; (2) delivers or is linked to providers that deliver primary health care; (3) coordinates a proactive health promotion program for students, faculty, and staff; and (4) relies on a collaborative interagency consortium (Killip et al. 1987). Examples of integrated school and community health programs include:

- ***School Health Curriculum Project (SHCP).*** As described in Table II.1, SHCP is a comprehensive school health education program that is designed to teach students how their bodies work by targeting discussion of a different body system (digestive, respiratory, cardiovascular, and central nervous) at each grade level. SHCP was developed by a coalition of the New York Academy of Medicine, the New York City Board of Education, and a coalition of foundations, corporations, individuals, and public and private agencies. SHCP initially was implemented in only five elementary schools, but is now offered in all New York City boroughs.

- ***School-Based Clinics.*** School-based clinics are comprehensive health centers that are located at or near schools and that provide a community-based, comprehensive approach to addressing the health needs of youths. Typically, school-based clinics involve the collaboration of schools and school boards, local health departments and community health centers, and the social service system.

These joint efforts of schools and communities, are for the most part, in the developmental and early implementation phases. As a result, there is no current evidence about their effectiveness in promoting school performance.

7. School Counseling

Children today can face a myriad of problems threatening their physical, intellectual, and social well-being. Suicide, divorce, teen pregnancy, substance abuse, child abuse and neglect, the threat of HIV/AIDS or other sexually transmitted diseases, and changing social and economic roles are only a subset of the complex problems confronting students. If the educational needs of students are to be satisfied, schools must consider many dimensions of student growth and development, including physical, intellectual, psychological, and social needs. School counselors figure prominently in holistic

and collaborative approaches to student development. They can provide preventive and curative services, using structured classroom instruction, individual or group counseling sessions, and parent/teacher consultations or in-service training (Thomas and Texidor 1987). Curriculum content may include assertiveness training; life-skills training; peer-led discussions; problem-solving training; and programs to enhance self-esteem, locus of control, and peer pressure resistance.

Although some counseling interventions cannot be identified clearly as health programs, others contain an obvious health component. One example of the latter is relaxation training. Teaching children relaxation skills has been shown to increase students' concentration and other "on task" behaviors in the classroom (Oldfield and Petoss 1986). Relaxation skills also have seen shown to reduce discipline problems, such as cutting or disrupting classes, fighting, and failing to do homework (Matthews 1986). Training in study skills and anxiety management has also been found to reduce test anxiety, increase self-esteem, and improve test performance (Wilson and Rotter 1986).

8. Health Promotion for School Faculty and Staff

Health problems of school faculty and staff parallel those of the general population. Common health concerns for this group include poor diet, inadequate levels of physical activity, hypertension, and smoking. Because school districts typically provide health care benefits to employees, they also shoulder the burden of medical claims resulting from morbidity and preventable mortality. In addition to medical costs, the academic and financial costs associated with absenteeism, disability, staff turnover, decreased productivity, and faculty and staff recruitment and replacement may be substantial.

Health promotion programs that positively affect the health behaviors of faculty and staff and reduce morbidity and premature death can provide substantial cost savings and academic benefits to school districts (Blair et al. 1987). Because of their existing infrastructure and professional resources, schools are the logical setting for work-site health promotion programs. Moreover, on-site health programs have been shown to reduce weight, body fat, blood pressure, anxiety, depression, and

smoking, and to increase physical activity levels. Participating teachers report a higher level of general well-being and a greater ability to handle job stress than do nonparticipating teachers (Blair et al. 1984). Faculty health promotion programs are also associated with reduced absenteeism, health care costs, and use of substitute teachers, as well as improved teacher morale and attitudes (Allensworth and Kolbe 1987). Faculty and staff who assume an active role in changing their behavior and improving their health may also serve as better models for students.

Health promotion programs for faculty and staff are not widely available, and well-designed evaluations of school-site programs are even more rare (Blair et al. 1987). In theory, improved morale and lower teacher absenteeism may translate into improved education outcomes for students; however, we are unaware of any evaluations that have assessed the impacts on students of faculty health promotion programs.

III. SCHOOL PERFORMANCE MEASURES

The first challenge in attempting to evaluate the impact of school health programs on school performance is to decide how to measure such performance. Typically, this is done by focusing on students' short-term and long-term education outcomes, behaviors, and attitudes. In addition, an evaluation must measure the characteristics of the students, their schools, and their communities that will shape school performance. Finally, an evaluation must develop measures to characterize the school health programs being evaluated.

Many alternative measures exist for each of these factors. A key to evaluation design is to select the measures most appropriate for detecting the expected impacts of the particular school health program under study. In this regard, the evaluation will consider such issues as the sensitivity of the measure to potential school health initiatives and the age range of students for which the measure is relevant.

In the following discussion of selecting measures to use in an evaluation of school health programs, we first examine the three components of school performance: (1) education outcomes; (2) student behaviors, and (3) student attitudes. We then consider the descriptive information that is needed to characterize the students, the schools and communities, and the school health intervention. Table III.1 provides an overview of specific measures that can be used to assess school performance.

A. MEASURES OF EDUCATION OUTCOMES

When researchers think about measuring student education outcomes, three questions come to mind. First, did students complete the educational process as expected? Second, did students learn what was expected of them? Third, what was the nature of the education process?

TABLE III.1

DATA REQUIREMENTS FOR MEASURING STUDENT SHORT-TERM EDUCATION PERFORMANCE

School Performance Measures	Descriptive Measures
Measures of Student Education Outcomes	**Student Demographic and Socioeconomic Characteristics**
Class Grades and Grade Point Average	Gender
Grade and Enrollment Information	Race
Grade Promotion	Age
High School Graduation Status	Ethnicity
Standardized Achievement Test Scores	Religion
Awards	Native Language
Course-Taking Patterns	Parents' Education
Courses Taken	Parents' Income
Classroom Instructional Level Grouping	Socioeconomic Status Measure
Classroom Placement	Number of Siblings
	Gender of Household Head
	Region of Residence
Measures of Student School Behaviors	**School Characteristics**
Attendance Information	Student/Teacher Ratio
Drop-Out Status	Expenditure per Student
Disciplinary Actions	Health Programs Offered
Time Spent Doing Homework	Percentage of Students Receiving Free Lunch
Extracurricular Activities	Drop-Out Rate
Employment and Earnings	Average Test Scores
	Socioeconomic Status of Students in School
Student School and Personal Attitudes	**Characteristics of Health Intervention**
Attitudes Toward Current School	Type of Health Intervention Received by School
Educational Expectations	Period in Which Health Intervention Was Offered
Occupational or Career Aspirations	Way in Which Intervention Was Implemented
Postsecondary School Plans	Type of Staff that Received the Intervention
Locus of Control	Degree of Student Participation in Intervention
Self-Esteem	

To measure whether students completed the education process as expected, researchers use such measures as:

- Whether students graduated from high school
- Whether students were promoted through the grade levels as expected
- Whether students currently are in a grade level that is appropriate to their age level

High school graduation status is an obvious way to measure whether students completed the education process. However, it is a long-term measure and is relevant only when evaluating outcomes for older students. Grade promotions provide a more detailed measure that can be used to evaluate outcomes for students of all ages. For short-term evaluations, the extent to which a class cohort moves from one grade to the next can be used to measure educational progress. For longer-term evaluations, either grade-promotion patterns spanning several years or a cohort's entire school experience can be used. When only cross-sectional data are available, prior grade-promotion patterns can be assessed by examining whether students are in the grades appropriate to their ages.

In assessing the second component of education outcomes--whether students learned what was expected--researchers use course grades, grade point averages, awards, and standardized achievement tests as measures. Course grades are the most basic indicator of whether a student learned what was expected from a course, and grade point averages provide a similar indicator for all courses taken during a given semester or set of years. Grades, however, present several disadvantages for evaluation purposes. First, they are a poor basis for making comparisons across schools. Grades reflect not only academic performance in specific classes, but also the curriculum of each school, the judgments of the teachers in that school about how well students learned, and the general levels of achievement within the school. The interactions of all of these determinants makes it difficult to use observed differences in course grades or grade point averages as a measure of differences in learning. Second, grades within a school may fail to detect overall increases in learning if they are awarded on

the basis of relative student performance (for example grading on a "curve"). The use of awards (such as honor roll status) to measure learning creates similar problems when attempting to make cross-school comparisons. In addition, such awards as the National Merit Scholar Awards tend to be given out relatively infrequently, making them poor indicators of the impacts of programs in specific schools.

Standardized achievement tests (such as the Iowa Test of Basic Skills or the California Achievement Tests) are better measures of academic achievement for making comparisons across schools. These tests assess how the knowledge of students taking the test compares with that of the broad samples of students used to develop the norms for the tests. Achievement tests may be specific, testing knowledge of a single subject, such as mathematics, or of a specific academic skill, such as reading level, or they may be quite broad, covering a wide range of topics.[1] Tests are also designed for each grade level. Table III.2 provides an overview of several prominent achievement tests, and of the grade levels for which specific tests are intended.

Although standardized achievement tests have advantages over grades as a measure of what students have learned, the tests have several disadvantages when being used for evaluation purposes. First, different school districts use different tests to assess their students. Thus, evaluations that compare schools in different districts must try to translate the results of the specific tests into a common basis. One approach to making tests comparable is to translate scores on individual tests into the corresponding percentiles. This approach was used in The National Assessment of Chapter 1 Services (U.S. Department of Education 1986). However, percentile rankings are not interval scales (that is, the intervals between percentile points do not represent equal levels of achievement, so that the achievement gain needed to improve from the 20th to the 30th percentile differs from the gain needed to move from the 40th to the 50th percentile). Thus, before average results could be compared across schools or districts, the percentiles were converted into Normal Curve

[1]Buros (1978) provides detailed explanations and examples of all of the major achievement tests given in the United States.

TABLE III 2

REPRESENTATIVE ACHIEVEMENT TESTS

	Grade Range													
Battery	K	1	2	3	4	5	6	7	8	9	10	11	12	13
California Achievement Tests	X	X	X	X	X	X	X	X	X	X	X	X	X	
Iowa Test of Basic Skills	X	X	X	X	X	X	X	X	X	X				
Iowa Tests of Educational Development										X	X	X	X	
Metropolitan Achievement Tests	X	X	X	X	X	X	X	X	X	X	X	X	X	
SRA Achievement Series	X	X	X	X	X	X	X	X	X	X	X	X	X	
Sequential Test of Educational Progress Series III (STEP III)[a]					X	X	X	X	X	X	X	X	X	
Stanford Achievement Test[a]		X	X	X	X	X	X	X	X	X				
Stanford Test of Academic Skills (TASK)									X	X	X	X	X	X
Tests of Achievement and Proficiency (TAP)										X	X	X	X	

SOURCE: *Psychological Testing*, 5th ed., 1982.

[a]CIRCUS extends STEP III to the preschool level; the Stanford Early School Achievement Test (SESAT) extends the Stanford series to the beginning of kindergarten.

Equivalents.[2] The problem for evaluation is determining whether translating test scores introduces additional noise into the measure, making it difficult to detect impacts, particularly those that are small.

Standardized tests also have weaknesses with respect to evaluating interventions at specific schools, because of potential mismatches between the tests and school curricula and because the tests may be too imprecise to measure the impacts of the intervention. Mismatches may result when the school curriculum and the tests emphasize different topics. In such cases, test scores may not reflect improvements in student learning at a specific school, because the test focuses on aspects of achievement that differ from students' actual achievement gains. Even when the standardized tests match a school's curriculum, the precision of the tests may be inadequate to measure a difference in achievement, unless large student samples are obtained. Thus, standardized tests would be inadequate for evaluations that compared, for example, classes of fifth graders in two schools.

B. MEASURES OF EDUCATION BEHAVIORS

The second aspect of school performance encompasses student school behaviors. These behaviors are one of the links between the health intervention and students' education outcomes. In addition, administrators consider many of the behaviors, particularly attendance, to be important program goals in and of themselves. In investigating impacts on school behavior, researchers generally ask the following questions:

- Are students currently attending school?
- Do they have any behavioral problems while in school?
- What is their degree of involvement in school?

---

[2]The Normal Curve Equivalent is one of many possible ways of using the normal distribution to convert percentile rankings into an interval scale that can be used to calculate averages. This method was used by Wang et al. (1981) to examine changes in levels of achievement in reading and mathematics.

Clearly, drop-out status is the most basic measure of a student's school attendance. Furthermore, because dropping out generally imposes a high price both on the student and on society, this measure has played a large role in evaluations of school programs (Rumberger 1987; Markey, 1988; and U.S. General Accounting Office 1987). This measure is appropriate for evaluations of school health interventions targeting high school students, and for evaluations of interventions affecting health behaviors that might influence a student to leave school. In particular, interventions designed to prevent teenage pregnancies or the spread of HIV infection might be expected to measurably affect drop-out status.

Attendance is another measure that is used to determine whether a student is in school. Typically, attendance is measured as the number of days in a semester that a student was not absent from school.[3] Some schools collect more detailed attendance information, indicating the number of classes missed. Attendance would be used to evaluate interventions focusing on health conditions that might cause a student to miss one or a few days of school, rather than to drop out. A school-based clinic is an example of the type of intervention that could be evaluated through the use of attendance data. By providing aspirin and a place to rest, a clinic may enable students who feel slightly ill to remain in school, thereby missing only one or two classes, rather than an entire day.

Information about patterns of absences, although typically not available, also would be useful for evaluation purposes. This information would help researchers to understand how a health program affected attendance patterns and would assist existing programs in implementing refinements to handle absence problems more effectively. For example, school-based efforts to reduce or prevent rates of serious illness would be expected to reduce the number of times that students were absent for a block of time, such as five or more consecutive class days. Alternatively, a school clinic might

[3]Schools differ to some extent about whether attendance information is maintained on a semester, quarter, or school-year basis. Evaluators typically aggregate attendance information to an annual basis in order to make comparisons across schools.

reduce rates of sporadic absences by enabling students with occasional headaches or minor nausea to remain at school, rather than having to be sent home.

After determining that students are in school, an important issue is whether they are disruptive. Researchers use measures of disciplinary actions to make this determination. The range of disciplinary actions encompasses relatively minor steps, such as after-school detention, to expulsions or suspensions. As an alternative to obtaining information on disciplinary actions, researcher might survey teachers about students' behavior in classes.

Finally, measures of antisocial behavior, including information about students' smoking, substance abuse, and contacts with the juvenile justice system, might be used. As discussed in the next chapter, however, the greatest obstacle to using these measures in evaluations of school programs is to obtain the data from the confidential records of the schools and juvenile justice system. In many cases, these institutions will not release such information, even when parents and students have given their informed consent.

To measure student involvement with school, researchers use such measures as time spent doing homework and participation in extracurricular activities. Information about the employment and earnings of students who work while enrolled in school also measures an aspect of school involvement.

C. STUDENT ATTITUDES

Student attitudes, the third element of school performance, comprise two broad categories: (1) school attitudes; and (2) personal attitudes. School attitudes include students' feelings about their current school, their current education expectations and aspirations, and expectations and aspirations

for postsecondary education. As one method of determining attitudes about school, surveys have asked students to respond by answering true or false to such statements as the following:[4]

- I am satisfied with the way my education is going.
- I have had disciplinary problems in school during the last year.
- I am interested in school.
- I have been suspended or put on probation in school.
- Every once in a while I cut a class.
- I don't feel safe at this school.

In addition, students may be asked to rank aspects of their current school, including the condition of the building, the quality of the instruction, teachers' interest in the students, the effectiveness and fairness of discipline, and the extent of school spirit. Questions to gain information on students' expectations tend to focus on such issues as how far the students expects to progress in school, and whether the student plans to continue his or her education after high school.

Personal attitudes pertain to an individual's self image, with the most common measures being locus of control and self-esteem. Other measures of personal attitudes are depression; attitudes about peers, attitudes about family, work and community; self-image, direction, control, and pride; and attitudes about the importance of marriage, children, money, and helping others.

Locus of control measures the degree to which an individual perceives himself or herself as having or lacking control over events (Rotter 1966). When an individual believes that an event has followed some action of his or her own, but that occurrence of the event is not entirely contingent on the action, it typically is perceived as resulting from luck, chance, or fate; as being under the control of powerful others; or as being unpredictable. An individual who interprets events in this manner exhibits an external locus of control. If the individual perceives the event to be contingent

[4]These questions were taken from the High School and Beyond Survey (University of Chicago, National Opinion Research Center 1980).

on his or her own behavior or personal characteristics, the person exhibits an internal locus of control.

Most locus of control studies use an abbreviated version of Rotter's Internal-External Control Scale (Rotter 1966). This scale contains such questions as:

- Do you think it's better to plan your life a good ways ahead, or would you say life is too much a matter of luck to plan ahead very far?
- When you make plans ahead, do you usually get to carry things out the way you expected or do things usually come up to make you change your plans?
- Some people feel they can run their lives pretty much the way they want to, others feel the problems of life are sometimes too big for them. Which one are you most like?

Locus of control appears to be linked to academic and occupational behaviors. Tesiny et al. (1980) studied schoolchildren and reported a significant positive relationship between external locus of control and depression, and a significant negative relationship between academic achievement and both external locus of control and depression. Andrisani and Nestel (1976) studied a representative national sample of nearly 3,000 adult men and found that internal locus of control is significantly related to indices of occupational success.

Self-esteem generally refers to individuals' positive or negative attitudes about themselves. Self-esteem is thought to be related to one's behavior and ability to function well in various roles. For example, a person with low self-esteem is more likely than one with high self-esteem to lack self-confidence, depend on others, be shy and guarded, value conformity, and lack creativity and flexibility (Wells and Marwell 1976). Typically, self-esteem is measured by asking students how often they believe the following statements to be true:[5]

- I feel that I am a useful person to have around
- I feel that I'm a person of worth

[5]These questions are derived from those in the Rosenberg self-esteem scale (Rosenberg 1965; Haggstrom et al. 1981; and Polit et al. 1985).

- I feel that I can't do anything right
- I feel that my life is not very useful
- I feel that I do not have much to be proud of
- As a person, I do a good job these days

D. DESCRIPTIVE INFORMATION

In addition to information about school performance, evaluations of school health programs will need to acquire basic descriptive information about the students who are eligible for the health program and about the schools in which the programs are implemented. Evaluations will also need information about the interventions themselves, particularly about which students received the intervention and the exact nature of the intervention. All of this descriptive information is used to describe the interventions under study and their operational context. As discussed in Chapter V, descriptive information is also needed in order to estimate program impacts. This information about students and schools is used to control for potential differences between groups of students that are due to factors other than the intervention under study.

An evaluation should try to measure all of the student characteristics that affect either the efficacy of the health program or school performance. In general, these characteristics include demographic information, such as age, gender, race and ethnicity, level of prior education, and whether a child is the first born. The list of characteristics also includes measures of household composition and socioeconomic status, including number of siblings, whether there is a single head of household, religion, household income level, education levels of the mother and father, how long the mother or father has worked, and how long the family has lived in the area. Finally, measures of family stress are used to control for the home environment of the student. These measures include such items as number of recent household moves and the number of parent separations or remarriages. Finally, such variables as the total amount of time spent in preschool are used to control

for additional education that might influence school performance (particularly of elementary school students).

The evaluation may also include variables about the student's health and overall access to health care before the intervention was instituted. Examples of these variables are: whether the student has any disability or substantial impairment; the number of days absent from school in prior years; information about insurance status; and whether the student has a regular health care provider.

Key characteristics about schools studied in the evaluation include measures of the intensity of education, such as student/teacher ratios, expenditures per pupil, school drop-out rates, average scores on achievement tests prior to the health intervention, and the percentage of students receiving free school lunches. Although more difficult to measure, it is also useful to obtain measures of school climate, that is, the extent to which the school offers a supportive environment that is conducive to learning. An index of school climate would include such factors as teachers' and students' assessments of school spirit and the adequacy of the facilities. It also should include an assessment of whether crime or substance abuse were felt to be important problems at the school.

Finally, the evaluation must acquire information about the health intervention under study, including information about the nature of the intervention and student participation. This descriptive information must be sufficiently in-depth to classify the intervention for evaluation purposes and to give users of the evaluation a basis for replicating the intervention. For classification purposes, researchers need to know how the intervention fits into the eight broad types of interventions described in Chapter II. In addition, more detailed information is useful, including information about how students obtain the intervention, its duration and intensity, and the content of the key program components.

Participation information indicates whether a student received the intervention, and how much was received. Measures of whether a student received a service are relatively straightforward for most programs, although obtaining such information may require extracting information from the

records of schools or specific school programs (such as the school lunch program or a school clinic). All students in a school typically are considered to be participants in school-wide programs (for example, an effort to alter the school environment). Measurement of the level of intervention will depend on the nature of the intervention. The extent of participation in health education programs would be measured by the number of class-hours or by whether students completed part or all of the curriculum. Participation in health service programs (such as a clinic) can be measured in general terms, such as number of visits, or in more precise terms that reflect the specific purposes of the visits (information, assessment, monitoring, or treatment) or specific nature of the visits.

IV. OBTAINING SCHOOL PERFORMANCE MEASURES

This chapter discusses the process of obtaining the data items that are needed to study the effects of school health interventions on measures of school performance. First, we discuss issues relating to the collection of school records data and describe additional data collection efforts to provide supplemental data. Second, we discuss the data from student surveys and school records that are contained in several existing national data sets. If information about the health programs in the schools covered by these surveys can be obtained, then data from these nationally representative samples of students can be used to determine the effect of health programs on student outcomes (see Chapter V).

A. SCHOOL RECORDS DATA

Most schools maintain records on individual students, which can readily be used to assess the effect of school health interventions on education outcomes. As institutional records, school records provide much detailed information and can be obtained at little burden to the schools themselves, while minimizing the recall or bias problems that are likely to arise when conducting surveys of students. Although the format, design, and content of transcripts vary across schools, the following core data items are common to most transcripts:

- Attendance (number of days absent from school)
- Grade promotion
- Drop-out status and length of stay in school
- Class grades and grade point average
- Course-taking patterns
- Disciplinary actions, suspensions, and expulsions
- Student's sex, race, and age

In addition, school records often contain standardized test scores.

These data items usually are stored at the school district, in computer files or in hard-copy form, although data on student disciplinary actions and drop-out status are often located in the files maintained at the individual schools. Thus, the process of obtaining student records involves collecting data both from school districts and from schools.

Local school files may also contain some or all of the following information: (1) teachers' commentaries; (2) internal memoranda concerning the student; (3) psychological reports; (4) results of diagnostic tests; and (5) reports of student criminal activity. However, because schools usually are reluctant to release these sensitive data items, this information typically cannot be used to study the effects of health interventions on school performance.

Obtaining transcripts on individual students is a three-step process. First, researchers who are evaluating a particular school program must gain the complete cooperation of the school districts that are part of the evaluation. Second, the researchers must obtain informed consent permitting the school to release student records. Third, the researchers must develop methods to collect the data.

1. Obtaining School District Cooperation

To evaluate successfully the effects of health interventions on school performance, it is imperative that officials from the school districts in the evaluation cooperate with program evaluators in all aspects of student record collection. Generally, it is not difficult to obtain the cooperation of officials from those districts whose schools are receiving funds to participate in a demonstration of which the evaluation is a component. Many school interventions are funded through federal or state grants, which often specify that recipient institutions must cooperate with designated program evaluators. Furthermore, the receipt of additional grants often depends on the results of the evaluation. Thus, grantee institutions, which have a stake in the outcome of the evaluation, have an incentive to cooperate with the evaluators.

It is usually more difficult, however, to gain the cooperation of school districts that do not receive funds to implement school interventions--that is, from institutions in a comparison (control group)--or the cooperation of districts conducting locally funded programs. Three methods typically are used to gain the cooperation of such school districts. First, financial incentives are offered, both as a means of partially offsetting costs to the districts of providing data and to recognize formally the districts' importance to an evaluation. Second, school district officials may receive aggregate information about student performance measures for students included in the evaluation. School district officials are often interested in obtaining this information, which enables comparisons to be made between the school performance of students in their districts and the performance of students in comparable school districts. Third, explaining to school district officials the importance of the evaluation to education policy often gains the cooperation of comparison-group schools.

2. Obtaining Informed Consent

The privacy and confidentiality of student data are always of concern to school districts. School districts face difficult conflicts between granting access to students for legitimate research purposes and maintaining parents' rights to approve or veto such participation. Thus, before releasing any data, school district officials must feel reassured that both their own policies and students' rights will be respected. To provide such reassurance, school districts are sent a statement of the evaluator's policy of and procedures for maintaining confidentiality. Most importantly, informed consent must be obtained from students in the evaluation sample.

In most cases, it is legal to obtain either active consent--that is, a signed, parental release form--or passive consent--that is, consent assumed to be given when a parent does not object to a publicized plan to collect data. To obtain active consent, school staff obtain from evaluators an informed-consent release form, the language of which authorizes the release of records data by the school. After the student samples have been selected, the staff are asked that students take the forms home, for signature by a parent. (A student who is of age can give his or her own active consent by signing

the form). A letter to the parents is attached to the consent form, explaining the demonstration, their own role in the evaluation, and the types of information that the researchers expect to obtain from school records.

When passive consent is used to obtain informed consent, schools are asked to use implied-consent procedures as a basis for releasing school records. Parents are sent a letter containing the information on the demonstration, but are asked to return a signed statement only if they choose to withhold permission for their child to participate in the evaluation. The legality of gaining access to student records through the use of passive consent falls under Title 20, Section 1232g, 6F of the Official Code of the United States, which states that educational institutions must release education records, even if parental consent has not been given, to:

> Organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or administering predictive tests, administering student aid programs, and improving instruction, if studies are conducted in such a manner as will not permit the personal identification of students and their parents by persons other than representatives of such organizations and such information will be destroyed when no longer needed for the purpose for which it is conducted.

The code suggests that, for authorized evaluations, the sponsoring federal agency may designate an organization to receive sensitive data, under the condition that the organization adopt certain methods of data protection as described by the code. The use of this code as a justification for passive consent has been approved by federal agencies for numerous school-based studies.

Our previous research suggests that school districts accept active consent more readily than they do passive consent. However, from a research perspective, the use of passive consent is generally preferable, because larger sample sizes can be obtained than when relying on active consent. For example, only about 20 percent of active-consent forms were returned to the schools in the School Nutrition Dietary Assessment Study (Burghardt et al. 1993). (When information on addresses and telephone numbers was obtained, and parents were contacted, however, response rates increased to 70 percent). By way of contrast, our experience suggests that, when passive consent is used, it is

possible to obtain a permission rate of 93 percent to 97 percent (in other words, only about 3 percent to 7 percent of parents withhold permission for their children to participate in an evaluation).

It is important to bear in mind that, even when active or passive consent is given, most school districts will not release some of the information contained in the student records and individual student files maintained in local schools, for fear that student rights will be violated. As discussed in the introduction to this section, protected information usually includes such items as reports of disciplinary actions related to criminal behavior and drug use, as well as diagnostic test results and psychological reports. Consequently, these measures generally cannot be used in evaluations.

3. Collecting Student Records Data

The first step physically in collecting student records data is to identify a contact person in each school district who is knowledgeable about the location and format of the records, and who has access to them. We have found that, even if some types of data are maintained at the school level, it is most efficient to collect school records data by working with a *single* data coordinator from each *school district*; under this type of arrangement, all data are assembled centrally and reported from a single source for each school district. The appropriate person to access and abstract information from the records is usually selected in consultation with officials from each school district.

The second step is to determine how the data are stored, and in what mode they are to be sent. Some school districts may have information that can be retrieved from centralized, routinely maintained data files. Other information may be available only from special records systems, which have been developed for the particular demonstration project, or from dispersed school records. Consequently, staff in some school districts will prefer to provide part or all of the requested data as computer file extracts (to disk or tape), whereas staff in other districts will prefer to provide the data as hard copy. Larger school districts, which are more likely than smaller districts to maintain centralized computer records, generally prefer to send computer extracts. Similarly, information about

high school students is more likely to be computerized than is information about elementary and middle school students.

The researchers must then determine which specific data items are to be sent, and in what format. To make these determinations, however, researchers must clearly recognize potentially competing concerns. On the one hand, the data items should be received in a well-defined, structured format. Such a format enables the data to be measured consistently across schools. It also obviates the costly necessity of having to recode data that have arrived in inconsistent or unstructured formats. On the other hand, the data collection process must acknowledge that different school districts use different data structures and maintain different types of records information. For example, although most transcripts are organized chronologically, some group courses by subject areas, or separate required courses from electives. Schools also use different types of grading systems (some schools do not even use a grading system). In addition, transcripts use a variety of formats to record information about failed courses, repeated courses, incompleted courses, and non-credit courses, such as remedial courses (some transcripts may not record these data). The types of data contained in the transcripts also vary. Some transcripts include test scores (for example, SAT scores, ACT scores, New York Regents exam scores, and the Iowa and California Achievement Test scores) and may even contain information about students' extracurricular activities. Furthermore, transcript information may be updated quarterly, biannually, or annually.

Our experience has shown that these competing concerns generally can be accommodated most efficiently by developing an abstraction form and working with the school districts in order to gain access to the school records required to complete that form. To design the form, we usually request a copy of the school records of several students (with all identifying information removed) and develop an abstraction form that facilitates the transfer of information from the record to the form. For school districts with automated school records data and with the capacity to send us computer data files, we request data items and file formats that are similar to those found in the abstraction

form. The form helps to ensure that all the required data items are clearly defined and taken into account when preparing a computer-file extract. Figure IV.1 displays a sample abstraction form that MPR sent to 20 school districts included in the Dropout Demonstration Assistance Program Evaluation to obtain student record information for approximately 3,000 students in 50 schools.[1]

Finally, when collecting student records data, it is useful to provide educational institutions with a transcript-request package that includes all materials that the schools will need to provide transcripts. This package includes a cover letter describing the study; a list of all students for whom transcripts are requested; a set of labels to be attached to the transcripts; signed active-consent forms (or the passive-consent forms); and forms that schools can insert into each student's record folder, indicating the student's participation in the study.

4. Obtaining Supplemental Data Items

School records data do not contain some of the data items that are needed to study the effects of school health interventions on school performance. To conduct a successful impact analysis, it is essential to obtain information about the health intervention itself. Furthermore, to improve the quality of the evaluation, it is important to collect supplemental measures of student demographic and socioeconomic status, and of school characteristics. Finally, teachers' assessments of students' behavior provide another important source of student-performance measures. Thus, additional data collection efforts should be considered in order to obtain these data items.

a. Obtaining Information on the Health Intervention

It is crucial to collect information about the health intervention itself--in particular, on whether a student receives the health intervention, on the degree of student participation, and on the characteristics of the intervention. These data items must be collected for three reasons. First, an

[1]For studies that include only small numbers of students in the evaluation sample, or small numbers of students per school district, it is sometimes more efficient to receive the raw records information, and to have the evaluator abstract the data items in-house.

FIGURE IV.1

Project:

DROPOUT DEMONSTRATION ASSISTANCE PROGRAM

STUDENT RECORDS FORM

Please fill in or correct missing/incorrect information on the left side for date of birth, Social Security Number, and student ID.

GRADES AND CREDITS

- Enter information in correct columns, based on number of grade reporting periods
- Enter "NE" if student was not enrolled in a particular period, "M" if information is missing from file, "NA" if not applicable.

	Check Reporting Periods Used 1 __ Semesters 2 __ Trimesters 3 __ Quarters 4 __ Summer														
	1 Circle One			2 Circle One			3 Circle One		4 Circle to Confirm Period		5 Attended Summer '92?	
	Semester 1	Trimester 1	Quarter 1	Semester 2	Trimester 2	Quarter 2	Trimester 3	Quarter 3	Quarter 4		01 __ Yes	00 __ No
a Grade Level		__			__		__		__		__	
b Grade Point Average		__.__			__.__		__.__		__.__		__.__	
c Total Credits Earned		__.__			__.__		__.__		__.__		__.__	
d Academic Credits Earned		__.__			__.__		__.__		__.__		__.__	
e Grade in Math	___ Numeric	or	___ Letter +/-	___ Numeric	or	___ Letter +/-	___ Numeric	or ___ Letter +/-	___ Numeric	or ___ Letter +/-	___ Numeric	or ___ Letter +/-
f Grade in English	___ Numeric	or	___ Letter +/-	___ Numeric	or	___ Letter +/-	___ Numeric	or ___ Letter +/-	___ Numeric	or ___ Letter +/-	___ Numeric	or ___ Letter +/-

NORM-REFERENCED TESTS

1 Test Name/Abbreviation	2 Name Code	3 Version/ Edition	4 Form	5 Level	6 Test Date	Score Type *(Check One)* 7 NCE or Percentile	Enter Score for Each Relevant Component 8 Composite Score	9 Reading Score	10 Math Score	11 Language Score	12 Other 13 Score / Subject Code	14 Other 15 Score / Subject Code
a	___	___	__	___	/ / Mo Day Yr	_ 01 or _ 02	___	___	___	___	___ / ___	___ / ___
b	___	___	__	___	/ / Mo Day Yr	_ 01 or _ 02	___	___	___	___	___ / ___	___ / ___
c	___	___	__	___	/ / Mo Day Yr	_ 01 or _ 02	___	___	___	___	___ / ___	___ / ___

BEST COPY AVAILABLE

FIGURE IV.1 (continued)

ATTENDANCE, DISCIPLINARY ACTION, AND ENROLLMENT STATUS

SCHOOLS ATTENDED

1 School _____ Code |_____|

a From: ___ / ___ / ___ (Mo / Day / Yr) b To: ___ / ___ / ___ (Mo / Day / Yr)

c Days Absent: 1 Excused |___|___|___|
2 Unexcused |___|___|___|
3 Total |___|___|___|

Check 'end reason' below *(Check One)*

01 | | End of Data Collection Period, Still In School
02 | | Dropped Out - Living In Service Area
03 | | Dropped Out - Moved Out of Service Area
04 | | Dropped Out - Whereabouts Unknown
05 | | Expelled
06 | | Transferred to Other School Within District/Service Area
07 | | Transferred to School Out of Service Area
08 | | Graduated from High School

Any suspensions in this period? From To

01 | | YES → a ___ / ___ / ___ (Mo / Day / Yr) b ___ / ___ / ___ (Mo / Day / Yr)
00 | | NO

c ___ / ___ / ___ (Mo / Day / Yr) d ___ / ___ / ___ (Mo / Day / Yr)

2 School _____ Code |_____|

a From: ___ / ___ / ___ (Mo / Day / Yr) b To: ___ / ___ / ___ (Mo / Day / Yr)

c Days Absent: 1 Excused |___|___|___|
2 Unexcused |___|___|___|
3 Total |___|___|___|

Check 'end reason' below *(Check One)*

01 | | End of Data Collection Period, Still In School
02 | | Dropped Out - Living In Service Area
03 | | Dropped Out - Moved Out of Service Area
04 | | Dropped Out - Whereabouts Unknown
05 | | Expelled
06 | | Transferred to Other School Within District/Service Area
07 | | Transferred to School Out of Service Area
08 | | Graduated from High School

Any suspensions in this period? From To

01 | | YES → a ___ / ___ / ___ (Mo / Day / Yr) b ___ / ___ / ___ (Mo / Day / Yr)
00 | | NO

c ___ / ___ / ___ (Mo / Day / Yr) d ___ / ___ / ___ (Mo / Day / Yr)

3 School _____ Code |_____|

a From: ___ / ___ / ___ (Mo / Day / Yr) b To: ___ / ___ / ___ (Mo / Day / Yr)

c Days Absent: 1 Excused |___|___|___|
2 Unexcused |___|___|___|
3 Total |___|___|___|

Check 'end reason' below *(Check One)*

01 | | End of Data Collection Period, Still In School
02 | | Dropped Out - Living In Service Area
03 | | Dropped Out - Moved Out of Service Area
04 | | Dropped Out - Whereabouts Unknown
05 | | Expelled
06 | | Transferred to Other School Within District/Service Area
07 | | Transferred to School Out of Service Area
08 | | Graduated from High School

Any suspensions in this period? From To

01 | | YES → a ___ / ___ / ___ (Mo / Day / Yr) b ___ / ___ / ___ (Mo / Day / Yr)
00 | | NO

c ___ / ___ / ___ (Mo / Day / Yr) d ___ / ___ / ___ (Mo / Day / Yr)

SOURCE: Evaluation of the Structured Demonstration Projects: Mathematica Policy Research, Inc.

BEST COPY AVAILABLE

impact analysis cannot be conducted unless the treatment group can be identified. Second, it is desirable to determine whether the health intervention has a greater impact on students who have a greater degree of involvement in the program than it does on those who are less involved. Third, it is important to identify the characteristics of the health intervention (for example, the types of services provided, how the program is administered, and the number and qualifications of program staff), so that successful and unsuccessful program components can be identified, and so that successful components can be replicated in other schools.

The particular process used to obtain data about a health intervention depends on the type of intervention being evaluated, because records on the health intervention may be maintained by service providers, local school districts, schools, teachers, and/or other units. For example, consider a study to evaluate the effects of a school-based clinic. Data on student visits to the clinic usually can be obtained from clinic records, and the characteristics of the clinic can be obtained by interviewing clinic staff. Now consider a study to evaluate the effects of a health education program that involves classroom instruction. In this case, school records data might provide information about the degree of student participation in the program, and the course content could be obtained by interviewing school administrators or teachers, or from course materials. As a third example, consider an evaluation of a school food service program. Information about the participation of a particular student, and about the specific foods eaten, could probably be obtained through student interviews. As a final example, consider a health program intended to change the school environment. In this case, it is most important to obtain information about overall student attendance at the school, because all attending students essentially participate in the program.

Finally, it is important to note that, if student records must be accessed, or if student surveys must be administered, it may be necessary to obtain active or passive consent. In addition, a process similar to the one used to collect student transcripts may be required in order to physically obtain the

information on the health intervention (for example, receiving copies of relevant data samples, developing abstraction forms).

b. Obtaining Information on Student Characteristics

Measures of student socioeconomic and demographic status (for example, family income, parental education, and household composition) are collected for an evaluation assessing the effects of a school health intervention on student performance measures. Measures of socioeconomic and demographic status often are used as matching variables, to select a comparison sample of students having average characteristics similar to those of the treatment sample (see Chapter V). The measures also are used to control statistically for any remaining differences in observable student characteristics between the comparison and treatment samples of students, when education outcomes between the two samples are compared. Finally, the measures can be used to determine whether impacts of the intervention differ by student socioeconomic and demographic subgroups.

Some of the data on demographic status are contained in student records. However, additional information on student demographic status and information on student socioeconomic status are usually obtained through student surveys, which can be administered in-person, by telephone, and/or by mail.[2,3] Informed consent is required before student surveys can be administered; the process of obtaining the consent is similar to the process used to obtain consent to collect school records data. It is important to realize, however, that student surveys are extremely costly to administer and code. Consequently, researchers must consider the net benefits to the evaluation of obtaining this type of survey data.

---

[2]The information can also be obtained through parent surveys, although this form of data collection is less frequently used than student surveys.

[3]Student surveys can also be used to obtain other data items that are not contained in school records data (see Chapter III). This information includes student educational and occupational expectations, student attitudes about school, time spent doing homework, student involvement in extracurricular activities, criminal activities, drug use, sexual activities, self-reported grades and attendance, and student employment and earnings.

c. Obtaining Information on School Characteristics

Information about the characteristics of schools (for example, student/teacher ratios, expenditures per student, and the percentage of students qualifying for free lunches) and information about student demographic and socioeconomic characteristics serve similar functions in evaluations of health interventions. The data items are used to ensure that the school environments of members of the treatment group and the comparison group are similar, and to control for any remaining differences between the school environments of the two groups when estimating program impacts. In addition, the information can be used to determine which subgroups of schools were affected most positively by the intervention.

From experience, we have found the Quality Education Database (QED) to be an excellent and cost-effective source of data on the school characteristics of 17,000 public school districts and 83,000 public schools at the elementary, middle, and high school levels. The QED contains the following data items:

- Names, addresses, and telephone numbers of the schools and school districts
- Student enrollment levels
- Racial and ethnic composition of the schools
- Number of teachers
- Number of schools in each district
- Discretionary dollars per pupil
- Percentage of college-bound students
- Personnel and job function codes

Additional data items can be obtained directly from school district records, although we have found the types of records maintained by different districts to be highly variable, as are the degree

of records automation and accessibility. The items that are contained in school district records, but not in the QED database, include some or all of the following:

- Average school standardized test scores
- School drop-out rate
- Percentage of students who qualify for free or reduced-price meals
- Percentage of students who are classified as limited or non-English proficient
- Disciplinary incidents
- Parental education

Experience suggests that the most efficient method of collecting data on these items is to send a standardized abstraction form to the school districts, requesting specific data items and formats.

d. Obtaining Teachers' Assessments

Teachers' assessments of students' academic performance and behavior are another important source of information about student school performance. Teacher-observation reports primarily provide qualitative information about student attitudes toward school, which are short-term outcomes likely to be affected by school health interventions. The following data items are a representative sample of the type of information that can be collected from teachers' assessments of students:

- Whether the student works hard to obtain good grades, and how often the student completes homework assignments
- Whether the student was recommended for academic honors or advanced placement classes
- How well the student relates to other students
- How often the student is absent or tardy
- How often the student is attentive in class
- Information on disruptive behavior

- Whether the student has an emotional or physical handicap
- Whether parent consultations on academic performance or behavioral or absentee problems were needed

Information on teachers' assessments of students typically is collected through teacher surveys, which are administered by mail. The first step is to produce a contact sheet for each student in the evaluation sample, listing the student's name, school, and other identifying information. These sheets, together with the teacher questionnaire, are sent to the data collection liaison in each school district. The liaison is instructed to distribute the contact sheets and questionnaires to the appropriate teachers; the teachers are asked to return the completed forms to the data liaison, who, in turn, returns them to the evaluator. Teachers usually are compensated for the time-consuming task of completing the questionnaires. Some information about teacher observations of students can also be obtained from student report cards and teacher memoranda, which are contained in the student files maintained by schools. However, as noted previously, these files are not readily released by schools, even after parents have signed consent forms.

B. NATIONAL DATA SETS

During the past two decades, the National Center for Education Statistics (NCES) of the U.S. Department of Education has funded several large-scale data collection efforts that obtained education data from students, teachers, and schools. The products of these data collection efforts potentially could be used to determine the effect of school health programs on school performance. The data collection efforts have identified aspects of school performance that are important to measure and also have developed survey questions that can successfully be asked of students. Thus, these questionnaires are prototypes that can be used in an evaluation of a school health intervention that involves collecting data from student surveys. In addition, it may be possible to use the data sets to determine the effect of health programs on education outcomes, if information about the health programs in the schools attended by the surveyed students can be obtained (see Chapter V).

The national data sets usually contain not only the education outcome measures from student records, but also the following data items from student and teacher questionnaires:

- Extracurricular activities
- Attitude toward current school
- Postsecondary plans
- Occupational or career aspirations
- Employment and earnings
- Measures of social competence, self-esteem, locus of control
- Student background and school characteristics
- Teacher-observation reports

Three national data sets have been widely used in the education literature: (1) the National Longitudinal Study of the High School Class of 1972 (NLS72) (Levinsohn et al. 1986; and Levinsohn and McAdams 1978); (2) High School and Beyond (HS&B) (Kolstad 1984); and (3) the National Educational Longitudinal Study of 1988 (NELS) (Ingels et al. 1976). Table IV.1 shows the major student performance measures from student surveys and transcripts contained in these data sets, which have the following key features:.

- **NLS72.** Administered by NCES, NLS72 is a nationally representative survey of more than 19,000 seniors from 1,061 high schools that began in the spring of 1972. Students were selected by using a two-stage probability sample, with schools as the first-stage units and students within schools as the second-stage units. At baseline, each sample member was asked to complete a student questionnaire and a test battery. School administrators were asked to supply survey data on each student, as well as student transcript data. Five followups, conducted in 1973, 1974, 1976, 1979, and 1986, were subsequently completed.

- **HS&B.** Also administered by NCES, the HS&B baseline survey was conducted in the spring of 1980. As with the NLS72 survey design, a two-stage probability sample was used to select students, with schools as the first-stage units and students within schools as the second-stage units. More than 1,000 secondary schools and more than 58,000 students are contained in the sample. HS&B cohorts include 10th graders and 12th graders. Follow-up interviews were completed in the spring of 1982, 1984, and 1986. The HS&B study did not collect transcript data on high school students.

TABLE IV.1

STUDENT SECONDARY SCHOOL EDUCATION PERFORMANCE MEASURES IN THE NLS72, HS&B, AND NELS DATABASES

Data Item	NLS72	HS&B			NELS		
		Sophomore Cohort		Senior Cohort			
		Baseline	1st Follow-Up	Baseline	Baseline (8th Grade)	1st Follow-Up (10th Grade)	2nd Follow-Up (12th Grade)
Measures of Student Education Outcomes							
Grade and Enrollment Information							
Grade level (S, T)	x	x	x	x	x	x	x
Type of high school program (S)	x	x	x	x	x	x	x
High school graduation status (T)	x		x	x		x	x
High school drop-out and transfer status (T)	x		x	x		x	x
Courses Taken							
Number of years of mathematics, English, history, languages, science, and vocational and technical courses in high school (S)	x	x	x	x		x	
Whether attended special high school programs (remedial, English as a second language, drop-out prevention, advanced placement) (S)	x	x	x	x		x	x
Courses taken in current school year, and opinions about the courses (S)					x	x	x
High school course titles, and number of credits earned (T)							x[a]
Standardized Test Scores							
Whether took SAT, ACT, or ASVAB tests (S, T)		x	x	x		x	
SAT scores (T)	x						x[a]
ACT scores (T)	x						x[a]
PSAT, Achievement Test, and AP scores (T)							x[a]
Whether passed minimum competency or proficiency test to obtain high school diploma (S)		x	x	x			x
Composite test score and quartile (S)		x	x	x			

BEST COPY AVAILABLE

TABLE IV.1 *(continued)*

Data Item	NLS72	HS&B			NELS		
		Sophomore Cohort		Senior Cohort			
		Baseline	1st Follow Up	Baseline	Baseline (8th Grade)	1st Follow-Up (10th Grade)	2nd Follow-Up (12th Grade)
Inconsistency score (S)	x						
Vocabulary formula score (S)	x						
Formula scores: picture numbers and mosaic comparisons (S)	x						
Course Grades							
Grades from 6th to 8th grade in English, mathematics, science, and social studies (S)					x		
High school grades in English, mathematics, science and social studies (S)						x	
High school course grades (T)							x[a]
Overall high school grades (S)	x						
Overall high school grades (T)		x	x	x			x[a]
Class rank (T)	x						x[a]
Awards							
Whether received any academic, athletic, community service, and/or attendance awards (S)						x	x
Measures of Student Education Behavior							
Attendance Information							
Number of days absent (S)		x	x	x	x	x	x
Days absent each year of high school (T)							x[a]
Main reason for last absence (S)						x	x
Number of days late (S)		x	x	x			
Number of classes cut (S)					x	x	x
Disciplinary Actions							
Whether sent to the office for misbehaving, not doing school work, or fighting (S)					x		
Whether parents received warnings about grades, attendance, or behavior (S)					x		
Whether suspended, put on probation, or arrested (S)						x	x

TABLE IV.1 *(continued)*

Data Item	NLS72	HS&B			NELS		
		Sophomore Cohort		Senior Cohort			
		Baseline	1st Follow-Up	Baseline	Baseline (8th Grade)	1st Follow-Up (10th Grade)	2nd Follow-Up (12th Grade)
Time Spent Doing Homework							
Average hours per week doing homework (S)	x	x	x	x			
Average hours per week doing homework in mathematics, science, English, social studies, and other subjects (S)					x	x	x
Extracurricular Activities							
Whether participated in athletics, clubs, youth groups, or organizations in and out of school during current school year (S)	x	x	x	x	x	x	x
Average amount of time spent reading, watching TV, socializing, etc. (S)		x	x	x		x	x
Employment and Earnings							
Whether currently/ever employed (S)		x	x	x	x	x	x
Current/most recent job: hours per week, type of work (S)		x	x	x	x	x	x
Current/most recent job: wages (S)		x	x	x		x	x
Hours per week spent on any job (S)	x						
Whether job is related to studies and will lead to future work (S)	x						
Measures of Student Attitudes							
Education Expectations							
Whether expects to graduate from high school (S)	x	x	x	x	x	x	x
Number of years until high school graduation (S)	x	x	x	x		x	x
Expected number of years in high school (S)					x		x
Whether expects to attend postsecondary school; if so, postsecondary school type (S)	x	x	x	x		x	x
Expected postsecondary field of study (S)	x	x	x	x			x
Reasons for choice of postsecondary school plans (S)	x						x
Whether expects to enlist in armed forces (S)	x	x	x	x			x
Whether expects to enter labor force after high school, and whether definite job is lined up (S)	x	x	x	x			x

TABLE IV.1 *(continued)*

Data Item	NLS72	HS&B			NELS		
		Sophomore Cohort		Senior Cohort			
		Baseline	1st Follow-Up	Baseline	Baseline (8th Grade)	1st Follow-Up (10th Grade)	2nd Follow-Up (12th Grade)
Occupational or Career Aspirations							
Expected type of jobs or occupations (S)	x	x	x	x	x	x	x
Reasons for choice of occupation (S)	x		x	x			
Attitudes About Current School							
Attitudes about school discipline, rules, facilities, safety, and spirit (S)		x	x	x	x	x	x
Attitudes about teacher/student relationships and grading (S)					x	x	x
Attitudes about school facilities, job placement record, academic instruction, vocational instruction, guidance and counseling, and teacher interest in students (S)	x						
Opinions about cheating, cutting classes, smoking, using drugs, fighting, talking back to teachers, disobeying school rules, etc. (S)						x	
Personal Outcomes	x	x	x				
Locus-of-control score (S)			x				
Self concept score (S)			x				
Work-, family-, and community-orientation scores (S)			x				
Whether has positive self image, direction, control, much to be proud of, etc.-(S)	x	x	x	x	x	x	x
Attitudes about importance of family, marriage, community, hard work, children, money, helping others, etc. (S)	x	x	x	x		x	x

*Contained in the NELS Student Transcript Survey and not yet publicly available.

NLS72 = National Longitudinal Study of the Class of 1972.
HS&B = High School and Beyond.
NELS = National Education Longitudinal Study of 1988.
S = Item contained in student surveys.
T = Item contained in student transcripts.
SAT = Scholastic Aptitude Test.
ACT = American College Test.
PSAT = Prescholastic Aptitude Test.
AP = Advance Placement.
ASVAB = Armed Services Vocational Aptitude Battery.

BEST COPY AVAILABLE

- **NELS.** Students in the NELS survey were also selected with a two-stage probability sample. More than 1,000 schools were sampled in the first stage, and almost 25,000 students were selected in the second. Unlike the two other surveys, however, NELS began with a survey of eighth graders in 1988. Follow-up interviews were completed in 1990 and 1992. The NELS data set contains student and teacher survey data, and high school transcript data.

These national data sets can be used to determine the effect of health programs on the education outcomes only of students between the 8th and 12th grades. A national data set does not currently exist that can be used to assess the effect of health programs on the education outcomes of primary school students. However, NCES plans to begin data collection for the Early Childhood Longitudinal Study (ECLS) in 1996. From 1996 to 2001, the ECLS will collect annual data on a nationally representative sample of approximately 25,000 kindergarten children. The study expects to obtain the following outcome measures:

- Academic achievement in reading and mathematics
- Attendance
- Grades
- Grade progression
- Special problem referrals: academic, behavioral, medical, speech/language
- Classroom instructional grouping level
- Classroom placement
- Child's expectations of school progress
- Child's disposition toward school
- Peer relations
- Extracurricular activities.

The ECLS also plans to collect data on school characteristics (for example, class size, student/teacher ratios, curricula), classroom practices, teacher characteristics, and school policies (for

example, discipline and attendance). At this time, the ECLS has no plans to collect information about school health programs. Nevertheless, as discussed in Chapter V, this proposed school-level data collection effort is an interesting evaluation design option for assessing the effects of school health programs on student education outcomes.

V. EVALUATION DESIGN OPTIONS

Evaluations of school health interventions can follow several alternative designs. In this chapter, we discuss the choice of evaluation design and present three major alternatives: (1) small-scale evaluations of local school health programs; (2) analysis of secondary data collected from nationally representative samples of students and prior school health program evaluations; (3) and multischool demonstrations.

Generally, program development proceeds through several stages, each of which has slightly different evaluation objectives. Initially, ideas for new programs or strategies emerge from the efforts of school districts and researchers to refine and improve the educational process. At this stage, the primary objectives of evaluations are to demonstrate the feasibility of the new ideas and to assess their potential to improve school performance. The evaluation should be simple, focusing on identifying promising strategies, rather than on generating precise impact estimates or developing estimates that will generalize to other schools. After the two primary objectives have been met, convincing evidence about the value of the intervention must be obtained. This information must meet stringent standards about precision and generalizability, in order to convince school administrators and school boards of the utility of the intervention.

Although school health interventions are well established, little is known about the types that have the greatest impact on school performance, or even about the general magnitude of the impacts. Given this lack of empirical evidence, the most appropriate evaluation strategies are those associated with the early feasibility stages of program development. Thus, it is appropriate to begin determining empirically which health interventions have the greatest impact on school performance by conducting evaluations of local school health evaluations. This strategy provides a means for systematically collecting information from the myriad of health programs that have been implemented in local schools, and to use that information to guide the development and evaluation of future programs.

It is also appropriate to identify potentially effective programs by designing an evaluation that exploits the large, existing databases. These databases often enable the relative effectiveness of alternative treatments to be assessed at relatively low cost.

As more is learned about school health strategies that affect school performance, more elaborate designs, such as multischool experiments, can provide the kind of rigorous test and precise information needed to support wide-scale program expansion. In some cases, such evaluations can be conducted by building on to existing evaluations that focus on the health outcomes of school health programs. In other cases, it will be necessary to initiate an entirely new demonstration and evaluation.

A. EVALUATIONS OF LOCAL SCHOOL HEALTH PROGRAMS

Sometimes, a school district will want to evaluate a school health program, but will lack the resources or sample sizes to pursue a large, rigorous evaluation. In these cases, school administrators must adopt an evaluation design that provides the most useful information, within budget constraints. This section describes some evaluation models that would be appropriate for such cases.

The distinguishing features of local evaluations are their focus on specific health issues, the need to consider both health outcomes and education outcomes, and their small overall scale. The need for these evaluations typically arises as local administrators consider how to modify the programs in their districts in order to address a specific problem, for example, a high drop-out rate among teenaged mothers, or disruptive behavior by students with mental health problems. The school district is also likely to be interested in the school performance of the students who have been targeted to receive the new services, as well as in the students' health outcomes. Finally, the school will often have the resources to implement the new intervention only on a small scale, in a single class or at a single school. Thus, the program may be loosely defined as a response to a specific problem, intended to achieve multiple objectives, and involving a relatively small number of students.

The primary audience for these evaluations is the school board, parents, teachers, and administrators within the district in which the change is contemplated. In addition, administrators who want to export their program ideas may be interested in convincing state or national policy makers, researchers, legislators, and education/health advocates about the desirability of the program change. Evaluations that are intended for local use only can adopt evaluation methods that will be sufficient for local policy makers to reach a decision. Evaluations whose audience is national must meet the more rigorous standards typically expected by such an audience.

In designing an evaluation that will meet the information needs of local or national audiences, the administrator must bear in mind the developmental stage of the program option and the magnitude of the change under consideration. For a new program, with many ill-defined details of program operations, the central issues are to refine the operational aspects of the program and to demonstrate the program's feasibility. In many ways, these issues correspond to the "hypothesis development" stage of a program; in that stage, administrators seek practical methods to solve problems relating to students' health and school performance.

The evaluation design that is appropriate for programs at an early stage of development differs from the design that is appropriate for well-defined programs that have been proven feasible, but that have not been compared with alternative programs having the same goals. For these well-defined programs, large, "hypothesis testing" evaluations, which focus on testing formal hypotheses about program impacts and cost-effectiveness, are appropriate. Thus, in a local school evaluation, large samples, extensive data collection plans, and rigorous designs are less appropriate than are efforts to describe the program and to document how services are delivered.

Small, less-precise evaluations may also be appropriate for local school initiatives because the consequences of adopting a specific program are often small. Program decisions can be reversed easily and may not involve substantial sums of money. In these cases, the costs of conducting an extensive evaluation would be disproportionate to the magnitude of the decision being made.

In general, carefully constructing a project description is the first step in evaluating a local program. This description provides the basis for understanding the program, as well as for making preliminary judgments about expected program effectiveness. The description must discuss the following characteristics:

- ***Program purpose***. What are the objectives of the new health program option, and why is the new program needed?
- ***Target population***. Who will the program target, and why is it appropriate that this group receive the new health services? (For example, what is the current health and education status of the anticipated target population?)
- ***Context***. What is the environment in which the new program will be fielded? (For example, what are the socioeconomic status of the community, size of the school population, richness of existing school and community health programs, and overall level of academic performance at the school?)
- ***Process***. How will the new program operate? (For example, how will students be selected to participate, what types of assessments will be performed in order to target the intervention to subpopulations, what is the curriculum, and how frequently will students be exposed to the intervention?)
- ***Expected outcomes***. What measurable aspects of health and school performance will the program affect, and how can these aspects be measured, using data available to the school?
- ***Rationale***. Why will this process help the anticipated target population to improve its outcomes within the current context of the school and community?

Information about all of these items can be collected with relative ease within a school district. Furthermore, a careful review of these descriptive factors can help administrators to judge whether an idea warrants actual implementation or continuation.

The next step in the evaluation is to describe the delivery of services. Specifically, the school district must document whether the services were delivered to students as planned, and whether the intervention actually increased service use. To make this determination, data must be obtained from school records about participation in the new program. The data might include the following items: the number of health classes provided; and the number of students attending; the number of students

tested; the number of brochures distributed; the number of counseling sessions held or the percentage of eligible students receiving a counseling session; the number of students receiving health care in the school; and the types of problems addressed.

When describing the services delivered, it is important to compare the extent to which the new program changed the level of intervention received by the target population with what would have happened in the absence of the new program (that is, the status quo). For example, relative to the lack of the program, did a school breakfast program increase the percentage of students eating breakfast? School data should be used to document the actual receipt of services. Schools will find it more difficult to assess the level of intervention that would have been provided under the status quo. In general, schools should try to make use of any easily obtainable data, such as information about the health status and school performance of the target population before implementation of the new program. Schools could also try to collect information from a comparison school as a way of estimating what would have happened in the absence of the new intervention. For example, to approximate the number of students who would have eaten breakfast in the absence of a new school breakfast program, students from a school without a breakfast program could be interviewed to determine how many ate breakfast. The estimates of changes in the receipt of services that are based on such a cross-school comparison are substantially more imprecise than are estimates that have been developed from large, pre-intervention/post-intervention multisite demonstrations. Nevertheless, the crude estimates are useful indicators of potential impacts, as long as the limitations of the methods are borne in mind.

It is essential that the school district document the costs of the health program. Cost data are one measure of the intensity of the intervention and are crucial to any effort to receive program approval from school boards or legislators. When measuring costs, it is often useful to estimate expenditures per student-month and the average duration of the intervention. Such estimates allow the administrator to distinguish between, say, the following two programs, each of which cost $300

per student. In one program, a $75-per-hour nurse might work intensively with a single student for four hours, to test for asymptomatic disease; in another, the nurse might make two, one-hour presentations per week for 40 weeks to a group of 20 students.[1]

The effort required to obtain this descriptive information will often exhaust the evaluation resources available to the district. For many purposes, however, this information is sufficient and would provide a reasonable basis for determining whether to continue or expand the new health program, or to return to the status quo. If time, interest, and resources permit, school districts could try to examine impacts over time, with the methods used to assess the extent to which the intervention changed exposure to the new services.

B. ANALYSIS OF EXISTING DATA

The purpose of this section is to describe designs that use existing data sources. In subsection B.1, we describe design options that use data from national surveys of students (see Chapter IV). We also discuss additional data collection efforts that are needed to evaluate the effects of the health interventions, and the statistical methods used to estimate program impacts. In subsection B.2, we present design options that use data from previous evaluations of the effects of school health interventions on health-related outcomes.

1. Analysis of Secondary Survey Data

Several large surveys of students have been conducted during the past two decades. Typically, these are surveys of nationally representative samples of students, which may involve student interviews, school records abstraction, teachers' observations, and school-level data collection. Integrating an effort to collect and analyze school health with an existing or planned large, cross-sectional survey of students is an interesting design option. With this design, each student would be

[1]The cost of the second option is $7.5 per student week (two hours of nurse time at $75 per hour, divided by the 20 students). The average student participates for 40 weeks. Thus, the average cost per student is $7.5 × 40, or $300.

linked with his or her school, information on the schools' health programs would be collected, and data on the programs would be linked with the student-level data.

It is important to emphasize that this design option does not rigorously evaluate the effects of health interventions on school performance. First, as discussed below, it may be difficult to estimate the impacts of the health programs, because nonexperimental statistical methods must be used in order to do so. Second, some of the surveyed schools may have offered a wide range of school health interventions, rather than one program, making it difficult to isolate the effects of a given health program from the combined effects of all of the programs offered. This blurring of effects defeats the primary purpose of the design option, which is to identify specific health interventions that seem to be correlated with improved school performance.

a. Basic Design

The three previously fielded surveys discussed in Chapter IV--the National Longitudinal Study of 1988 (NELS), High School and Beyond (HS&B), and the National Longitudinal Study of the High School Class of 1972 (NLS72)--offer the opportunity to identify school health program that may improve school performance. The basic design entails the following four steps: (1) identifying the schools attended by the students in the data set; (2) developing protocols for contacting the schools; (3) collecting retrospective data on the school health programs operating when the survey was conducted; and (4) obtaining measures of student participation in the programs.[2]

The major limitation of using existing surveys to identify potentially promising school health interventions is the difficulty of the fourth step--obtaining accurate information about the school health programs used by the surveyed students. It will be difficult or impossible to identify an intervention to evaluate if the health program records are missing or incomplete, if it is difficult to locate and interview program administrators, or if program administrators do not remember details

[2]It may be feasible to collect information only on a subset of the schools in the data sets, because about 1,000 schools were sampled for each fielded survey.

about the characteristics of the health programs.[3] Furthermore, it may be difficult and costly to determine the degree of student participation in the interventions, if the only way to obtain this information is to interview the students.

A more feasible design for collecting school-level data on health programs from a large cross-sectional or longitudinal survey of students is in the context of surveys that are *currently* being designed. In particular, the Early Childhood Longitudinal Survey (ECLS) is a key candidate for this evaluation design strategy. As described in Chapter IV, the ECLS is a longitudinal survey focusing on school readiness, children's transition to school, and children's progress through the elementary grades. The survey has two components:

1. A large-scale, birth-cohort survey of 10,000 to 15,000 children
2. A school-based survey of approximately 25,000 kindergarten children

The kindergarten cohort is of primary importance for this design report, because pre-kindergarten school health interventions although desirable, are both less prevalent and less clearly articulated. However, the general approach to evaluating preschool programs would resemble that for the kindergarten cohort.

For each child in the sample, the ECLS will collect data on the child, the child's family, community characteristics, school characteristics, and education outcomes. Table V.1 presents preliminary specifications of variables to be collected by the ECLS. Of particular interest is the column on school characteristics and programs. To assess school health programs, one would add a set of variables to this column about the school health programs implemented in each school. As discussed here, this design would involve the development of a survey module to collect information from school staff about school health programs.

---

[3]These limitations should be less severe if the NELS database, rather than the HS&B or NLS72 database, is used. Because the NELS database has been fielded more recently, it would be easier to obtain information about school health programs and student participation from this database.

TABLE VI

PRELIMINARY SPECIFICATIONS OF VARIABLES TO BE COLLECTED BY THE EARLY CHILDHOOD LONGITUDINAL SURVEY

Child	Family	Community	School/Child Care Program	School Outcomes
Gender Race Age Language Socio-Emotional Maturity Level (Self Concept, Self-Esteem, Adaptive Behaviors) Cognitive Developmental Level (General Knowledge; Verbal, Nonverbal, Cognitive Style) Physical Developmental Level (Fine and Gross Motor Skills) Health and Well-Being Birth History	Parent Characteristics • SES/income • Education • Maternal employment status • Language • Attitudes, beliefs, and values (e.g., toward learning, self-efficacy, school) • Expectations for self/child • School/community involvement Family Characteristics • Household/family composition (no. of adults in home, no. of older/younger children in home) • Mobility • Health and well-being Family Processes • Home activities • Child rearing practices Family Resources • Extended family	Community Characteristics • Size/density • Geographical location (urbanicity) • Income level • Ethnic diversity Institutional Community Resources Used (e.g., Libraries, Museums, Parks, Parent Outreach Programs, Churches)	Classroom/Child Care Setting • Group/class size • Student/teacher ratio • Curriculum/program focus • Subject matter schedule Classroom Practices • Instructional level grouping • Content materials (manipulatives, paper/pencil) • Teaching style • Family involvement activities Teacher Characteristics • Expectations/attitudes (e.g., self-efficacy, control) • Gender • Age • Race • Training/education level • Years of experience Physical Setting • Size • Location (urbanicity) • Ethnic/income diversity • Resource programs • Community outreach programs • Affiliation (public/private) • Cost • Staff mobility • Length of school day School Policies • Discipline • Attendance • Physical safety • Family involvement activities	Academic Achievement • Reading • Mathematics • Subject matter achievement Attendance Record Grades Grade Progression Special Problem Referrals • Academic • Behavioral/discipline • Medical • Protective service • Speech/language Classroom Placement • Regular classroom • Non-regular (e.g., resource, self contained, day treatment) Child's Expectations of School Progress Child's Attitude Toward School Peer Relations Extracurricular Activities

SOURCE: U.S. Department of Education, National Center for Educational Statistics, *Draft Proposal for a Longitudinal Study Program of Young Children*, May 1992.

SES = socioeconomic status.

BEST COPY AVAILABLE

b. Definition of the School Health Program

To survey schools about the types of school health interventions operating at the time of the national surveys, it is first necessary to specify the information that will be used to describe these programs. Programs can be described broadly according to the eight general categories discussed in Chapter II.[4] More refined measures can also be developed that would distinguish, for example, multi-semester and single-semester health education curricula, or school-based and school-linked clinics. The more refined measures might also identify the specific goals of each program, thereby classifying programs in terms of goals and approach. Allensworth (1993) developed one system for classifying program goals. In this system, goal categories include: (1) increasing knowledge about health; (2) promoting wellness and health maintenance; (3) developing students' decision-making skills and sense of responsibility for their health; and (4) integrating physical, mental, emotional, and social dimensions of health.

In addition, it is essential to note instances in which multiple types of health programs were implemented, particularly such promising combinations as health education and school clinics. Finally, information about specific program features should be collected, such as information about the curriculum for a school health class or the eligibility criteria and content of a school breakfast or lunch program. This detailed information will enable researchers to develop more detailed program classifications, and researchers and administrators to describe more fully any programs that look promising on the basis of the exploratory analysis of the survey data. Program features that are important to document include program intensity (which can be measured by expenditures per pupil or student:teacher ratios) and program content (which includes such variables as the number of days or hours per week that a physical education program is required, the types of services offered to

[4]The eight types are: (1) school health education; (2) school health services; (3) school environmental improvement programs; (4) school food service; (5) physical education and fitness; (6) integrated school and community health promotion programs; (7) school counseling; and (8) health promotion for school faculty and staff.

students through a school-based or school-linked clinic, and the curriculum content of a substance abuse prevention program).

Ideally, the school health data could be linked to individual students, to identify exposure to specific health programs at each grade. It is also critical to determine whether the student participated in any school-wide or voluntary program. If a school offers a counseling program, the evaluation would need data on both the content of the program and individual students' use of the services of a school counselor. When content and participation data are obtained, the analysis could examine the effects of both *program availability* and *program participation* on school performance.

Although a wide range of health programs potentially can be evaluated, it may be difficult to obtain reliable estimates of the impacts of a particular type of program, if it is offered by only a small number of schools, or if it is often offered in conjunction with other health programs. Consequently, the analysis might focus only on health programs that were offered in isolation by many schools. It might also be desirable to identify schools that did not offer a given program, but that had characteristics similar to those of the schools that did. The mean school performance measures of students in the treatment sample could then be compared with the corresponding measures of students in the comparison sample. This approach might reduce some problems due to sample selection bias, which are discussed in the next subsection.

c. Methodological Approach

The basic objective of the evaluation design is to compare the relevant education outcomes of students exposed to a specific school health program or receiving specific health interventions with the outcomes in the absence of the exposure. It is not possible to make such a comparison precisely, because the outcomes of students under those two regimens cannot be observed. However, strong evaluation designs attempt to achieve this objective by comparing performance measures of students receiving specific health interventions with performance measures of a comparison group of students

who did not receive the interventions.[5] Under this evaluation design, students in schools with a defined set of school health programs will be compared with comparable students in schools that provide different health programs or, possibly, no health programs.

An important issue is the endogeneity of the school health variables. The evaluation must account for whether schools that offer specific health programs differ systematically from schools that offer either other health programs or no well-defined health program at all. If these groups of schools differ systematically (other than with respect to the types of health programs offered), then between-school differences in school performance may reflect the other systematic differences, rather than the effect of the specific health program being evaluated. For example, if wealthier school districts are more likely than poorer districts to offer a planned, sequential, comprehensive health program, then estimating the effects of this type of program on the students' education outcomes must control for the income level of the school district. If this is not done, education outcomes that might result from the differential income level would be attributed to the school health program.

Regression models offer one means of controlling for systematic differences between schools that offer a specific health program and schools that do not, *when all such differences can be measured*. More complex procedures are needed to correct for the sample selection bias that could arise when some of the between-school differences cannot be measured. In addition, the appropriate estimation process will depend on whether a school health program is mandatory for eligible students, or is voluntary. To illustrate these statistical models, first consider the case that is simplest from a

[5]The fundamental method used to select a comparable group is random assignment. As discussed in the following section, evaluations based on random assignment have an overriding advantage over evaluations with alternative designs: observed differences in outcomes between groups that are randomly assigned to treatment and control status can be attributed to program impacts with a known degree of statistical confidence. Random assignment is not feasible, however, with the evaluation design option of analyzing secondary survey data collected from nationally representative samples of students.

statistical perspective: a mandatory program in which all systematic differences between schools can be measured.[6] In this case, the regression model would take the mathematical form:

$$(1) \quad Y_{is} = \alpha + X_i'\beta + Z_s'\gamma + \delta H_s + \epsilon_{is}$$

where:

Y_{is} = a school performance measure for student i in school s

X_i = a vector of characteristics for student i

Z_s = a vector of characteristics for school s

H_s = an indicator variable that equals 1 if school s offers a particular mandatory health program, and equals zero if school s does not

α = the intercept term

β = the vector of regression coefficients for the student characteristic variables

γ = the vector of regression coefficients for the school characteristic variables

δ = the measure of the impact for the health intervention

ε_{is} = a random disturbance term for student i in school s and represents unobserved factors that affect Y_{is}

In the absence of an unmeasured systematic difference between schools, ordinary least squares techniques can be used to obtain unbiased estimates of the impact of the health program on a particular measure of school performance. This impact will be represented by δ (if the dependent variable is continuous). In this case, it is assumed that, after controlling for the X and Z variables, there are no remaining systematic differences between students who received the health intervention

[6]We define a mandatory health program as one requiring participation of all eligible students.

and those who did not. By including additional program-specific indicator variables, this approach can be extended in a straightforward fashion to address differences across several health programs.[7]

As noted, although this approach assumes no remaining systematic differences between students who received the health intervention and those who did not (after controlling for observable characteristics), the decisions of schools to offer a particular health program may be based on unobserved or unmeasured school characteristics that also affect the outcome of interest. In this case, the standard regression approach will produce biased estimates, because the health program indicator variable (or variables) will be correlated with the error term in the outcome equations.

We can attempt to control for such unobserved factors by applying the two-stage econometric technique developed by Heckman (1976, 1979). This method first models and estimates a school's decision to offer the evaluated mandatory health intervention. It then includes in the regression equations variables that are functions of the estimated probabilities of receiving the intervention.

In the first stage, a model is specified to explain how schools decide to offer a health program. This decision may be determined by school staff, the school board, and/or state mandated programs. In the model, IN_s is an index of the net benefits to the decision makers from schools offering a particular health program and is a linear function of observed variables W_s (some of which may also be in Z_s), and unobserved variables u_s. Then:

$$(2) \quad IN_s = W_s'\theta + u_s$$

In terms of this function, the health program will be adopted if the index function exceeds zero, and will not be adopted otherwise. Parameter estimates in equation (2) are obtained by using probit

[7]The equation could also be made more flexible with a set of interactions between the health program dummies and other explanatory variables; however, for expositional purposes, we have not done so.

binary choice statistical procedures, and are then used to obtain the predicted probability that each school offers the health program.

Heckman corrects and tests for sample selection bias by using a function of the predicted probability of offering a health program (known as Mills ratio) to obtain an estimate of the portion of the error term that is correlated with H_s, and to include the Mills ratio as an additional regressor in the outcome equations, thereby eliminating the correlation between H_s and the error term. Because the remaining error term is no longer correlated with H_s, the ordinary least squares estimates of α, β, γ, and δ will be consistent (unbiased in large samples).[8]

This technique is theoretically sound. However, its success in producing reliable impact estimates depends on successfully identifying the outcome equations, which take the form of equation (1). The outcome equation is identified successfully if it is possible to distinguish between the coefficients of the outcome equation and the coefficients in the equations that predict whether a school offers the health program. The usual identification method is to specify variables that affect the decision to offer the health program, but that do not affect education outcomes (that is, at least one variable in W_s must be excluded from Z_s in equations (1) and (2)). In practice, such variables are often difficult to specify. Thus, whether these sample selection bias models can be estimated successfully will depend on whether such variables can be found and on the health outcome being evaluated. An example of such a variable in our context is a dummy variable indicating whether a health program is state mandated. This variable is likely to be correlated with whether or not a student receives the health intervention, but is not likely to be correlated with the students' school performance.

---

[8]The two-stage econometric technique developed by Maddala and Lee (1976) is another commonly used method to control for sample selection bias. This approach also uses a probit model to obtain predicted probabilities that a school offers the evaluated health program. However, rather than including an estimated mills ratio term for students in the outcome equations, the approach replaces the indicator variable H_s with the predicted offer probabilities. These predicted probabilities are uncorrelated (in large samples) with the error term in the outcome equation ϵ_s (if the control variables W_s are uncorrelated with ϵ_s), and consistent impact estimates can be obtained by using ordinary least squares techniques.

The statistical procedures discussed thus far can also be used to obtain consistent impact estimates for ***voluntary*** health programs, although there are now *two* sources of potential sample selection bias. First, selection bias may result from unobserved systematic differences between the characteristics of schools that offer a particular health program and the characteristics of those that do not. Second, selection bias may result from differences between the students who chose to participate in the programs and those who did not.

For example, consider the following regression equation:

$$(4) \quad Y_{is} = \alpha + X_i'\beta + Z_s'\gamma + \eta H_s + \pi H_s * P_i + \epsilon_{is},$$

where:

P_i = an indicator variable of the health program participation status of student i that equals 1 if the student participated in the program, and zero otherwise

η = the measure of the effect of the health program on nonparticipants who attended schools that offered the health program relative to students in schools that did not offer the health program

π = the measure of the effect of the health program on participants relative to nonparticipants in schools that offered the health program,

and where Y_{is}, X_i, Z_s, H_s, α, β, γ, and ϵ_{is} are defined as in equation (1). In this example, selection bias can occur if: (1) H_s is correlated with the error term ϵ_{is}; and/or (2) P_i is correlated with ϵ_{is}. Thus, to obtain consistent impact estimates, it is necessary to control for *both* possible sources of selection bias. The two-stage statistical procedures involve modeling and estimating a school's decision to offer the evaluated health intervention and the student's decision to participate in the program, and then including in the regression equations variables that are functions of the first-stage predicted probabilities.

2. Analysis of Data from Previous Evaluations of School Health Interventions

An interesting and cost-effective design option is to use data from the numerous demonstrations and evaluations that have studied the effects of school health programs on student *health-related*

outcomes, such as health knowledge, attitudes, behavior, or actual health status (see Fasciano and Devaney 1993, for a review of these studies). This option is attractive, because these studies have already collected most of the data items needed to evaluate the effects of health programs on education outcomes, including: (1) the definitions and characteristics of the health programs being evaluated; (2) the degree of student involvement in the programs; and (3) some characteristics of the students and schools in the evaluation samples. In addition, the data can be readily accessed and analyzed, because they have been coded and data-entered.

If retrospective school records data for students in the evaluation samples can be obtained, the study data can be used to evaluate education effects of school health interventions. These school records would be obtained through the relevant school districts. Negotiations with school district officials would be needed in order to determine what informed-consent documents (if any) would be required for release of school records.[9]

However, school districts may not keep student records for an extended period. Thus, school records for the students in some of the older studies might no longer exist. Furthermore, the databases from the previous studies are unlikely to contain identifier variables, such as student names or social security numbers, that would simplify the process of linking the study data with the student data. Unless the researchers who conducted the previous studies can provide identifier variables for linking with school records, this evaluation method cannot be used.

If school records can be obtained for students in a particular study, one must then decide on the methodological approach to determine the impacts of the health program on school performance. The choice will depend on the type of sample design that the study used. For example, if the study collected data only on program participants, it would not be possible to determine how the relevant

[9]It would be costly to obtain active consent from students. First, it would be necessary to locate the students, so that consent forms could be sent. Second, the students would need to be contacted by telephone, to ensure that a high percentage of active consent forms were returned. It would be less costly to obtain passive consent, because follow-up calls would be unnecessary.

outcomes of students exposed to the health program would compare with their outcomes in the absence of the program. It would be possible to determine only whether the program had larger effects on students who had greater exposure to the program than it did on students with lesser exposure. However, if the study had been a large demonstration, in which only some students participated, then the appropriate methodological approach is identical to the approach discussed in Section C.

Finally, it is important to decide which of the existing evaluation databases should be used for this design option. We offer four criteria. First, if the original evaluation showed that the evaluated school health intervention had significant impacts on health outcomes, then that school health intervention would appear to be a good candidate for affecting school performance. The model presented in Chapter II suggests that school health programs directly affect health outcomes, and that, in turn, these health outcomes affect school performance. Thus, it is not likely that a health intervention would affect school performance, but would not affect health-related outcomes. Second, it must be possible to obtain school records data for the students in the database. Third, the database should contain clear, detailed information about both the characteristics of the health programs and student participation in the programs. Fourth, the study must have used a sensible sample design. For example, it is not desirable to use data from a study that employed too few students or schools to make reliable inferences, or whose treatment-group members had different characteristics from comparison-group members.

Several studies that might be expanded to include education outcome measures are currently planned or under way. One such study is the proposed follow-up to the Child and Adolescent Trial for Cardiovascular Health (CATCH). Funded by the National Heart, Lung, and Blood Institute (NHLBI), CATCH is evaluating the effectiveness of a three-year elementary school program designed to discourage smoking and promote physical activity and healthful eating. If funded by NHLBI, the follow-up study will gather post-intervention data on the students when they enter sixth grade, and

again when they enter eighth grade.[10] NHLBI also is funding an evaluation of an obesity prevention program, modeled on CATCH, for Native American students. The researchers will begin protocol development in the fall of 1993; thus, school performance indicators could be included among the study's primary outcome measures. The Centers for Disease Control and the National Institute for Nursing Research also are funding several large-scale studies that might be expanded to explore the effects of health interventions on education outcomes.

C. LARGE-SCALE, MULTISITE DEMONSTRATIONS

The purpose of this section is to discuss experimental and comparison-group designs to conduct rigorous evaluations of the effect of school-based health interventions on student school performance. The core of the proposed design methodologies is the comparison of a treatment group of students, that receives a health intervention, with a control (comparison) group, that does not. In an experiment, schools or students are randomly assigned to treatment status. In a comparison-group design, treatment-group members are matched with a group of students who have similar characteristics, but who do not receive the intervention (or who receive an alternative intervention). In a properly designed evaluation, the comparison group represents what would have happened to the treatment group, had the latter not received the intervention.

In this chapter, we consider designs in which treatment-group members attend schools that offer the health intervention, and control-group members attend schools that do not. We adopt this approach for two reasons. First, a key objective of most school health programs is the school-wide modification of norms and behavior. Second, if members of both the treatment and control groups attend the same school, crossover effects may produce inaccurate impact estimates. Crossover effects

---

[10]School records data on students' baseline school performance would have to be collected retrospectively for this study. Although this collection process is straightforward logistically, intensive negotiations with the schools may be needed to obtain permission to collect the data. This would be true especially for schools that felt they had previously reached a final agreement about their need to provide information to the CATCH Study.

occur when control-group students are exposed to the treatment through their interactions with students in the treatment classrooms.

In the rest of this section, we discuss the superiority of randomized designs over comparison-group designs. We then discuss, for mandatory and voluntary health programs, design options to evaluate the effect of *new* school health programs on school performance, and sample size requirements. We also discuss design options to evaluate the effect of *existing* school health programs and conclude with the importance of performing process analyses.

1. The Case for Random Assignment

Two concepts central to sound evaluation design are: (1) internal validity; and (2) external validity. Internal validity relates to whether what we observe--for example, a measured difference in student attendance rates--is in fact caused by the intervention. External validity pertains to whether observed demonstration impacts would be replicated if the intervention were implemented in broader settings, or on a larger scale.

In terms of internal validity, well-designed randomized experiments are clearly superiority to comparison-group methodologies. Only with random assignment do we have a basis for attributing what we observe to the impact of the health intervention with a known degree of statistical precision. The classical statistical methodology underlying randomized experiments requires that we compare two independent, random samples--both drawn from the same population, but only one receiving the intervention. When two such samples are compared, simple statistical tests tell us whether any observed differences are due to chance. If the tests indicate that the differences are unlikely to be due to chance, the differences can be attributed directly to the intervention. Random assignment creates the conditions for applying this classical methodology, unless the randomization process or method of introducing the intervention creates contaminating effects, which could be confused with the impact of the intervention.

With comparison-group procedures, assumptions and statistical models must eliminate differences between the samples that could result from sources other than the intervention (see the discussion in Section B). If these efforts are successful, remaining differences can be attributed to the intervention, possibly with some measure of statistical confidence. However, if sources of unmeasured or unmeasurable difference exist, this approach produces biased impact estimates. These problems are commonly termed ***sample selection bias*** in the statistical literature and were discussed in Section A of this chapter.

Due to concerns about conducting methodologically rigorous evaluations, most recent evaluations of school health programs involve the use of random assignment (see Fasciano and Devaney 1993). Typically, entire schools are randomly assigned to treatment or control groups.

2. Sample Design for New School Health Interventions

The purpose of this subsection is to discuss design options to answer the following question:

- What is the effect of a proposed health intervention that does not currently exist in schools on students' school performance?

The design options are discussed separately for mandatory health programs, such as health education programs, and for voluntary programs, such as school-based clinics.

a. Mandatory Health Programs

When all targeted students are required to participate in a health program, the ideal experimental design to evaluate the impact of the program would randomly assign a nationally representative sample of *students* (not schools) to treatment or control status. Student records data and student questionnaire data (if the budget permits) would be collected, with a comparison of mean school performance measures representing the impact of the intervention. These impact estimates would have internal and external validity.

However, randomly assigning students, rather than schools, to treatment or control status typically is not an operational design option. Most health interventions are either school-based or are presented in a classroom setting. Thus, providing a health intervention to a small number of treatment students in a given school would be extremely cost-ineffective and would disrupt the normal operations of the school. Furthermore, it would be costly to collect student records and questionnaire data for small numbers of students in a large number of schools.

A more cost-effective option is a two-stage, clustered design. In the first stage, schools are randomly assigned to treatment or control status. Treatment schools would be required to provide the health intervention to targeted students; control schools would not. In the second stage, students are randomly selected from the treatment and control schools, and school performance measures are collected for this sample.[11]

A major limitation of the two-stage, clustered design is that, when the number of schools is small, random assignment does not ensure initial comparability between treatment and control schools. Generally, in such cases, the evaluation designs are associated with low statistical power. A potential solution to this problem is to use a stratified random sampling plan, which matches, prior to random assignment, groups ("blocks") of schools on the basis of variables that are highly correlated with outcome variables. The procedure identifies a grouping factor, creates groups of schools that differ in terms of that factor, and randomly assigns schools to treatment and control groups from within each group. Using this plan increases the degree of initial comparability between the treatment and control schools and increases the power of the evaluation design.

A second, related limitation of this design option is that many schools must be sampled in order to detect program impacts of a reasonable size. This problem occurs because the variances of the

[11]An alternative two-stage, clustered sampling design option is to randomly choose schools to participate in the demonstration in the first stage, and to randomly allocate students *within* each school to a treatment or control group in the second stage. However, as discussed in the introduction to this section, we do not consider this design option to be feasible.

impact estimates have a school-specific component that can be very large when only a few schools are sampled.

An important issue to address in conducting school health evaluations from a two-stage, clustered design pertains to the unit of assignment versus the unit of analysis. There are two schools of thought on this issue. One holds that the units of assignment and analysis should be the same, and that the appropriate level of analysis is the smallest group for which variation in the program is provided, which is generally the school (or classroom). Other researchers use the school or classroom as the unit of assignment and the student as the unit of analysis. Their counterargument rests on three basic points. First, student-level analysis is necessary to gauge the effects of programs than span more than a school year, since students do not always move to a new grade in intact classes nor do they always stay in the same school. Second, many of the most important predictors of certain health behaviors, such as substance abuse, are individual characteristics; student-level analysis allows more precise controls for preprogram differences among groups. Third, student-level analysis allows researchers to ascertain whether different types of students respond differently to a program.

The level of assignment does not matter for the analysis if there are no school-specific components to the variance. In this case, the impact of the health intervention is estimated by comparing mean school performance measures *per treatment-group student* with mean school performance measures *per control-group student*, or by comparing mean performance measures *per treatment school and control school*. For example, consider two treatment and two control schools, and that ten students are sampled from each school. Assume that the students in the two treatment schools were absent during a given semester for a total of 80 days, and that students in the two control schools were absent for a total of 100 days. The impact analysis can be performed either by comparing mean treatment-school and control-school absentee rates (40 days versus 50 days), or by comparing mean student absentee rates (4 days versus 5 days). Mean outcomes per school are simply a multiple of mean outcomes per student.

However, in the more typical case, at least some portion of the impact variance will reflect a school-specific component. In this case, using the student as the unit of analysis when larger units are used for assignment ignores the correlation among students' knowledge, attitudes, and behavior within a given school and violates the statistical assumption of the independence of the error term (Campanelli et al. 1989). The solution to this problem is to adjust for the effect of the clustering of students within schools. The "design effect" correction, calculated from the between- and within-group correlations for each dependent variable, indicates the factor by which simple random sample variance calculations must be increased to correct for clustering. Without this correction, student-level analysis tends to overestimate the statistical significance of results. With this correction, however, it is legitimate to use students as the unit of analysis when schools or classrooms are the unit of assignment. We now discuss this issue in greater detail.

b. Statistical Precision

Classical statistical theory applies equally to samples of students and samples of schools. Determining sample sizes and the appropriate number of schools from which to draw the sample can be considered a two-stage sampling problem, in which both the number of students and the number of schools contribute to the variance of relevant impact measures. The variance of the mean value of a school performance measure for *students* in treatment schools can be written as:

$$(5) \quad Var\ (\mu) = \frac{(1-\rho)\sigma^2}{n} + \frac{\rho\sigma^2}{a}\left(1-\frac{a}{A}\right)$$

where:

- μ = mean value of the school performance measure
- σ^2 = variance of the school performance measure
- ρ = between-school variance as a proportion of total variance of the school performance measure
- a = number of treatment schools in the sample

A = total number of schools in the relevant universe

n = number of students from treatment schools in the sample[12]

The variance of the mean value of a school performance measure for students in control schools can be written similarly, and the variance of the impact estimate is the sum of variances of the mean values for the treatment and control samples.

Several features of the variance calculation should be noted:

- The first term of the variance measure is the conventional expression for the variance of a mean value (σ^2 divided by the total sample size), reduced by the $(1-\rho)$ term.
- The second term is the between-school component of variance, deflated by the number of schools in the universe. It also includes a "finite population correction" term reflecting the proportion of all schools included in the sample.[13]
- As the proportion (a/A) approaches 100 percent, the variance for the clustered design approaches the variance of a randomized design.
- The between-school variance typically constitutes a small proportion of the total variance, but can vary widely by performance measure. In our experience, the value of ρ lies between .01 and .03 (that is, between 1 percent and 3 percent). Although this component is small, it is deflated by the *number of schools* in the sample, rather than by the number of observations. Thus, it becomes a major portion of the variance of the mean impact measure.

We now discuss the minimum detectable health program impacts on various student school performance measures for simple random-sample and two-stage, clustered designs, and for treatment and control samples of various sizes. We assume that the sampling universe is a school district that contains 40 schools (A = 40, in equation (5)).[14]

---

[12]The use of this formula assumes the same number of students in each school in the relevant universe.

[13]There is also a finite population correction for student samples, but it can be safely ignored, except when sampling rates are very high.

[14]The use of the school district as the sampling universe is commonly used in the education literature (see for example, the Evaluation of Structured Demonstration Projects, which is currently being performed by MPR). The calculations are made on the basis of statistical tests that compare mean outcomes of treatment-group members with those of control-group members.

Table V.2 shows minimum detectable impacts on student absentee and drop-out rates when a simple random sample of students (*not schools*) is assigned to either a treatment or control group.[15] The table shows that, for a sample of 200 treatment-group students and 200 control-group students, we would expect to detect a significant reduction in the percentage of the school year absent if the *true* program impact were 3 percentage points or more. This reduction corresponds to roughly 5 school days, assuming a typical school year of 180 days. Similarly, we would expect to detect a significant reduction in drop-out rates, if the true program impact were 12.4 percentage points or more. If the total sample size increases five times, to 2,000 students, the minimum detectable impacts are reduced by more than 50 percent.

Table V.3 shows minimum detectable impacts on several student outcome measures for a two-stage, clustered design.[16] The minimum detectable impacts are presented for different numbers of treatment and control schools sampled in the first stage, and for different numbers of students sampled in the second stage. The calculations assume the between-school variance is 2 percent of the total variance of the outcome measures.[17]

Three main points emerge from Table V.3. First, for a given number of students in the sample, the minimum detectable impacts for the two-stage, clustered design are usually much larger than they are for the simple random-sample design. For example, the minimum detectable impact for drop-out rates is 13 percentage points when 1,000 students and 10 schools are included in the sample. However, the corresponding figure for a simple random sample of 1,000 students is only 7.9 percentage points (see Table V.2). Second, for a given number of students in the sample, statistical power can be improved by increasing the number of schools in the sample. For example, if the

---

[15]The calculations are based on statistical tests that compare mean outcomes of treatment- and control-group members.

[16]Again, the calculations are based on statistical tests that compare mean outcomes of students in the two groups.

[17]In our experience, the between-school variance is typically 1 percent to 3 percent of the total variance.

TABLE V.2

MINIMUM DETECTABLE IMPACTS ON VARIOUS EDUCATION PERFORMANCE MEASURES, USING A RANDOM-SAMPLE DESIGN

Treatment/Control Sample	Percent of School Year Absent (Mean = .12, SD = .12)[a]	Drop-Out Rate (Mean = .50, SD = .50)
100/100	4.2	17.6
200/200	3.0	12.4
500/500	1.9	7.9
1,000/1,000	1.3	5.6

NOTE: All calculations assume a 95 percent one-tailed test for significance and 80 percent power.

[a]The mean and standard deviation for the percent of school year absent are from the Evaluation of Summer Training and Employment Program (Sipe et al. 1988)

TABLE V.3

MINIMUM DETECTABLE IMPACTS ON VARIOUS EDUCATION PERFORMANCE MEASURES, USING A TWO-STAGE, CLUSTERED DESIGN

Treatment/Control Schools	Treatment/Control Students	Percent of School Year Absent (Mean = .12, SD = .12)[a]	Drop-Out Rate (Mean = .50, SD = .50)
2/2	100/100	5.9	24.4
	200/200	5.1	21.1
	500/500	4.5	18.8
	1,000/1,000	4.3	18.0
5/5	100/100	4.9	20.3
	200/200	3.9	16.1
	500/500	3.1	13.0
	1,000/1,000	2.8	11.8
10/10	100/100	4.5	18.7
	200/200	3.4	14.1
	500/500	2.5	10.3
	1,000/1,000	2.1	8.8

NOTE: All calculations assume a 95 percent one-tailed test for significance, 80 percent power, and that the between-school variance is 2 percent of the total variance.

[a]The mean and standard deviation for the percent of school year absent are from the Evaluation of Summer Training and Employment Program (Sipe et al. 1988)

sample contains 2,000 students, the minimum detectable impact for drop-out rates decreases by more than 50 percent when the number of schools in the sample is increased from 4 to 20. Third, statistical power is increased more rapidly by increasing the number of schools in the sample than by increasing the number of students in the sample. For example, the detectable impact for drop-out rates is about the same whether 200 students are sampled from 20 randomly chosen schools (4,000 students in total), or whether 2,000 students are sampled from 4 randomly chosen schools (8,000 students in total).

As this example demonstrates, to obtain internally valid impact estimates, sample sizes must be larger in two-stage, clustered designs than in simple random-sample designs. The analysis also suggests that, if relatively large student samples are specified within schools, statistical power can be improved primarily by reallocating the sample in order to increase the number of schools. The appropriate number of schools and students depends on the particular health intervention being evaluated. However, the sample should be large enough to reliably detect effects that are large enough to justify implementing the health program.

c. Voluntary Health Programs

The two-stage, clustered design can also be used to evaluate the effect of a voluntary school health program on school performance. In this case, *two* types of program impacts can be estimated: (1) impacts for students who participated in the program; and (2) impacts for *all* students in the treatment schools, including nonparticipants.

The techniques to obtain impact estimates for all students in the treatment schools are identical to the techniques used to obtain impact estimates for mandatory health interventions.[18] However, because students who choose to participate in the health program comprise a self-selected sample that may, therefore, systematically differ in observable and unobservable ways from students who do

[18]It may be difficult to detect any impact if only a small number of students receive the intervention (unless there were substantial crossover effects in the treatment schools).

not choose to participate, different techniques must be used to obtain impact estimates only for those who *received* the treatment. Thus, impact estimates generally *cannot* be obtained by comparing the mean performance measures of program participants with the mean performance measures of students in the control schools; students in the latter group include both those who would choose to participate in the health program, if it were offered, as well as those who would not.

Unbiased estimates for these impacts can be obtained by comparing the mean values of the performance measures of all students in the treatment schools with those of the control students. The impact estimate for a particular school performance measure is given by the following equation:

$$(6) \quad \textit{Impact} = \frac{\bar{y}_t - \bar{y}_c}{p}$$

where:

$\bar{y}_t$ = mean value of the school performance measure for *all* students in the treatment schools

$\bar{y}_c$ = mean value of the school performance measure for students in the control schools

p = proportion of students in the treatment schools who participate in the health program[19]

This impact estimate has lower statistical power than does the impact estimate for mandatory health programs, because one must account for the variance of $(\bar{y}_t - \bar{y}_c)$, as well as for the variance of the proportion p, when calculating the variance of the impact estimate in (6). Thus, when evaluating voluntary health programs, the number of sampled schools and students must be greater than when evaluating mandatory health programs, to achieve comparable minimum detectable impacts for those who received the intervention.

---

[19]This formula assumes that students in the treatment schools who chose not to participate were not systematically influenced by students who chose to participate.

3. Sample Design for Existing School Health Interventions

The purpose of this subsection is to discuss design options to answer the following question:

- What is the effect on student school performance of a mandatory or voluntary health intervention that currently exists in schools ?

In general, randomized designs cannot be used to conduct rigorous evaluations of existing health programs, because the treatment was not randomly allocated to schools or students. Schools that chose to implement the health program are a self-selected sample and are likely to have observed and unobserved characteristics that systematically differ from schools that did not implement the program. Thus, comparison-group designs must be used to estimate the program impacts of existing health interventions.

a. Mandatory Health Programs

An external comparison-group design can be used to assess the effect of an existing mandatory health program. The comparison sample comprises students who attend schools that do not offer the health program and who meet the general profile of the eligible students in the treatment schools. Comparing mean values of the performance measures for treatment and control students yields estimates of program estimates; regression analysis can adjust for *measured* between-group differences.

This approach is severely limited by possible sample selection bias resulting from *unobserved* differences between the participant and external-comparison groups.[20] Earlier in this chapter, we discussed statistical methods to control and test for selection bias using cross-section data. As noted, developing effective selection bias models generally depends on finding at least one explanatory

[20]Studies by Lalonde (1986), Lalonde and Maynard (1987), and Fraker and Maynard (1987) illustrate the limitations of statistical matching techniques, even when used in conjunction with regression analysis. In their evaluations of youth samples, there was strong evidence of selection bias in the impact estimates after matching.

variable that affects whether a student receives the treatment, but that does not affect the outcome. In practice, it is very difficult to find such variables. Second, even effective models can substantially increase impact standard errors (by a factor of 3 to 6, from our experience) relative to the impact standard errors obtained from a randomized design that has the same sample size. Thus, the sample size requirements to achieve a given level of statistical precision are much larger in non-randomized than in randomized designs.[21]

b. Voluntary Health Programs

External comparison-group designs can be used to evaluate the effect of voluntary health programs on school performance. The procedure involves matching students who participate in the health program with students having similar measured characteristics who attend schools that do not offer the program. *Both* participating students and nonparticipating students in the treatment schools can also be matched with students in the comparison schools. Impact estimates are then obtained by using the methods discussed previously.

An alternative strategy is to obtain internal comparison groups. This approach involves selecting a comparison group from among eligible students who choose not to participate in the intervention. With this approach, internally valid impact estimates can be obtained when using performance measures for students in the *same* school. Because the between-school variance component is not included in the variance calculations for the impact estimates, statistical power may be greater than when the external comparison-group method is used. However, sample selection bias is likely to be

[21]Methods to control for sample selection bias can also be used with longitudinal data. These methods involve comparing the **change** in school performance measures of students in the treatment group before and after the intervention with the corresponding change in school performance measures of students in the comparison group. This "difference-in-difference" estimator controls for sample selection bias, because unobserved student characteristics are removed when differences in student outcomes are taken. However, pre-program data on school performance must be available, and comparing education outcomes across different grades must make sense. For example, scores of standardized tests taken in the 8th grade are probably not comparable with those from the 12th grade.

greater, because students who choose to participate in the health program are likely to differ from nonparticipating students in important and, in some cases, unmeasured ways relating to the outcome of interest. Although regression analysis can correct for measured between-group differences, it is difficult to control for unobserved between-group differences when using models that correct for sample selection bias.

4. Process Analysis

Thus far, we have considered quantitative methods to evaluate the impact of a health program on school performance. However, an important component of a demonstration evaluation is a process analysis that documents fully how the program operates. A process analysis is crucial for understanding the intervention being tested, and for providing the contextual information to support and enhance interpretation of findings from the impact evaluation. If, for instance, the impact analysis indicates that a particular health program substantially affects school attendance rates, then the process analysis will help to document how the program operated when those impacts were obtained, and what services were provided. Information obtained in the process analysis also helps to understand why a health program had only limited impacts (or no impacts), or why the program was found to be effective for some subgroups of participants, but not for others.

Although the focus of the process analysis of a health intervention will depend on the type of intervention that is being evaluated, the following basic issues should be addressed:

- Program implementation (who provided the services, what specific services were provided, where services were provided, and how long services lasted)
- The organizational characteristics of the health projects and the schools in which they are being administered
- The necessary conditions for replicating successful programs, including formal and informal links between the program being evaluated and other health or education programs

- The process by which students were selected to receive services, including any screening or assessment instruments used to identify health needs
- The costs associated with the health programs

Four primary forms of data collection specifically designed to support a process analysis are: (1) staff interviews; (2) on-site observation; (3) focus groups; and (4) records review. Staff interviews and questionnaires provide information about the characteristics and background of project staff, their roles in delivering services, and their assessments of the success of the health programs. On-site interviews and observation provide important information about program operations, organization, and implementation. Focus groups provide insights into the intangible factors that can lead to program success, and the factors that affect students' responses to the health programs. Finally, program records provide a basis for estimating costs and tracking the flow of students through the program.

REFERENCES

Anastasi, Anne. *Psychological Testing*. 5th ed. New York: MacMillan, 1982.

Andrisani, Paul J., and Gilbert Nestel. "Internal-External Control as Contributor to and Outcome of Work Experience." *Journal of Applied Psychology*, vol. 61, no. 2, 1976, pp. 156-165.

Allensworth, Diane D. "Health Education: State of the Art." *Journal of School Health*, vol. 63, no. 1, January 1993, pp. 14-20.

Allensworth, Diane D., and Lloyd J. Kolbe. "The Comprehensive School Health Program: Exploring an Expanded Concept." *Journal of School Health*, vol. 57, no. 10, December 1987, pp. 409-412.

Anderson, C.L., and W.H. Creswell. *School Health Practice*. St. Louis, MO: C.V. Mosby Company, 1976, pp. 94-96.

Blair, Steven N., Thomas R. Collingwood, Roger Reynolds, Mike Smith, R. Donald Hagan, and Charles L. Sterling. "Health Promotion for Educators: Impact on Health Behaviors, Satisfaction, and General Well-Being." *American Journal of Public Health*, vol. 74, 1984, pp. 147-149.

Blair, Steven N., Leonard Tritsch, and Suzanne Kutsch. "Worksite Health Promotion for School Faculty and Staff." *Journal of School Health*, vol. 57, no. 10, December 1987, pp. 469-473.

Brink, Susan G., and Philip R. Nader. "Comprehensive Health Screening in Elementary Schools: An Outcome Evaluation." *Journal of School Health*, vol. 54, no. 2, February 1984, pp. 73-78.

Burghardt, John, Nancy Chapman, Barbara Devaney, Tom Fraker, Philip Gleason, and Anne Gordon. "School Nutrition Dietary Assessment Study." Princeton, NJ: Mathematica Policy Research, Inc., forthcoming.

Buros O.K. (ed.). *The Eighth Mental Measurement Yearbook*. Lincoln, NE: University of Nebraska Buros Institute of Mental Measurements, 1978.

Campanelli, Pamela C., T.E. Dielman, Jean T. Shope, Amy T. Butchart, and Deborah S. Renner. "Pretest and Treatment Effects in an Elementary School-Based Alcohol Misuse Prevention Program." *Health Education Quarterly*, vol. 16, no. 1, spring 1989, pp. 113-130.

Chubb, John E., and Terry M. Moe. "Politics, Markets, and America's Schools. Washington, DC: The Brookings Institution, 1990.

Clayton, R.R. "The Delinquency and Drug Use Relationship Among Adolescents: A Critical Review." In *Drug Abuse and the American Adolescent*, edited by D.J. Lettieri and J. Ludford. NIDA Research Monograph 31. Rockville, MD: U.S. Department of Health and Human Services, 1981.

Dryfoos, Joy. "School-Based Health Clinics: A New Approach to Preventing Adolescent Pregnancy?" *Family Planning Perspectives*, vol. 17, no. 3, March/April 1985, pp. 70-75.

Dryfoos, Joy, and Lorraine Klerman. "School-Based Clinics: Their Role in Helping Students Meet the 1990 Objectives." Report to Carnegie Corporation, July 1987.

Duncan, Burris, W. Thomas Boyce, Robert Itami, and Nancy Puffenbarger. "A Controlled Trial of a Physical Fitness Program for Fifth Grade Students." *Journal of School Health*, vol. 53, no. 8, October 1983, pp. 467-471.

Ellickson, Phyllis L., Robert M. Bell, and Kimberly McGuigan. "Preventing Adolescent Drug Use: Long-Term Results of a Junior High Program." *American Journal of Public Health*, vol. 83, no. 6, 1993, pp. 856-861.

Eggert, Leona L, Christine D. Seyl, and Liela J. Nicholas. "Effects of a School-Based Prevention Program for Potential High School Dropouts and Drug Abusers." *The International Journal of the Addictions*, vol. 25, no. 7, 1990, pp. 773-801.

Fasciano, Nancy, and Barbara Devaney. "A Study to Design an Evaluation of the Effects of School-Based Health Interventions on School Performance: A Review of the Literature on School-Based Health Interventions." Princeton, NJ: Mathematica Policy Research, Inc., February 1993.

Flay, Brian R., David Koepke, Shirley J. Thomson, Susanne Santi, J. Allan Best, and K. Stephen Brown. "Six-Year Follow-Up of the First Waterloo School Smoking Prevention Trial." *American Journal of Public Health*, vol. 79, no. 10, 1989, pp. 1371-1376.

Fraker, Thomas, and Rebecca Maynard. "The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programs." *Journal of Human Resources*, vol. 22, no. 2, 1987, pp. 194-227.

Gilliam, Thomas B., Susan E. MacConnie, David L. Geenen, Albert E. Pels, and Patty S. Freedson. "Exercise Programs for Children: A Way to Prevent." *The Physician and Sports Medicine*, vol. 10, no. 9, September 1982, pp. 96-108.

Haggstrom, Gus W., Thomas J. Blaschke, David E. Kanouse, William Lisowski, and Peter A. Morrison. *Teenage Parents: Their Ambitions and Attainments*. Santa Monica, CA: RAND Corporation, 1981.

Hawkins, J. David, and Richard F. Catalano. "Broadening the Vision of Education: Schools as Health Promoting Environments." *Journal of School Health*, vol. 60, no. 4, April 1990, pp. 178-181.

Heckman, James J. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variable and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement*, vol. 5, 1976, pp. 475-492.

Heckman, James J., and Thomas E. MaCurdy. "Labor Econometrics." In *Handbook of Econometrics*, edited by Zvi Griliches and M.D. Intriligator. New York: North-Holland Publishing Company, 1983.

Ingels, Steven J., Leslie A. Scott, Judith T. Lindmark, Martin R. Frankel, Sharon Myers, and Shi-Chang Wu. *National Education Longitudinal Study of 1988, First Follow-Up: Student Component Data File User's Manual*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, April 1992.

Killip, Diane C., Sharon R. Lovick, Leslie Goldman, and Diane D. Allensworth. "Integrated School and Community Programs." *Journal of School Health*, vol. 57, no. 10, December 1987, pp. 437-444.

Kirby, Douglas, Cynthia S. Waszak, and Julie Ziegler. "An Assessment of Six School-Based Clinics: Services, Impact and Potential." Washington, DC: Center for Population Options, October 1989.

Kolbe, L. "Indicators for Planning and Monitoring School Health Programs. In *Proceedings of the Symposium on Indicators of Health Promotive Behaviors*, edited by S. Kar, E. Berkanovic, and S. Churgin. Los Angeles, CA: University of California, Los Angeles Press, 1986a.

Kolbe, L. "Increasing the Impact of School Health Promotion Programs: Emerging Research Perspectives." *Health Education*, vol. 17, no. 5, 1986b, pp. 47-52

Kolstad, Andrew. *High School and Beyond: A National Longitudinal Study for the 1980s. Data File User's Manual*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, April 1984.

Kominski, Robert. "Estimating the National High School Dropout Rate." *Demography*, vol. 27, no. 2, May 1990, pp. 303-311.

Kuntzleman, Charles T., and Debra A. Drake. "The Feeling Good Youth Fitness Report." Report submitted to the W. K. Kellogg Foundation. Spring Arbor, MI: Fitness Finders, Inc., 1984.

Lalonde, Robert J. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review*, vol. 76, no. 4, 1986, pp. 604-620.

Lalonde, Robert J., and Rebecca Maynard. "How Precise are Evaluations of Employment and Training Programs: Evidence from a Field Experiment." *Evaluation Review*, vol. 11, no. 4, 1987, pp. 428-451.

Levinsohn, J.R., L. Lewis, J.A. Riccobono, and R.P. Moore. *National Longitudinal Study Base-Year, First, and Second Follow-Up Data File User's Manual*. Research Triangle Park, NC: Research Triangle Institute, 1976.

Levinsohn, J.R., and K.C. McAdams. *The National Longitudinal Study Data Base Design: A User Survey*. Research Triangle Park, NC: Research Triangle Institute, 1978.

Lohrmann, David K., Robert S. Gold, and Wanda H. Jubb. "School Health Education: A Foundation for School Health Programs." *Journal of School Health*, vol. 57, no. 10, December 1987, pp. 420-425.

Madalla, G.S. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, MA: Cambridge University Press, 1983.

Madalla, G.S., and Lung-Fei Lee. "Recursive Models with Qualitative Endogenous Variables." *Annals of Economic and Social Measurement*, vol. 5, 1976, pp. 525-545.

Markey, J. "The Labor Market Problems of Today's High School Dropouts." *Monthly Labor Review*, vol. 111, no. 6, 1988, pp. 36-43.

Matthews, D. "Discipline: Can It Be Improved with Relaxation Training?" *Elementary School Guidance and Counseling*, vol. 20, no. 3, 1986, pp. 194-199.

McGinnis, J.M., and Christopher DeGraw. "Healthy Schools 2000: Creating Partnerships for the Decade." *Journal of School Health*, vol. 61, no. 7, September 1991, pp. 292-297.

McKenzie, J.F., and I.C. Williams, "Are Your Students Learning in a Safe Environment?" *Journal of School Health*, vol. 52, no. 5, 1982, pp. 284-285.

Meyers, Alan F., A. Sampson, M. Weitzman, B. Rogers, and H. Kayne. "School Breakfast Program and School Performance." *American Journal of Diseases*, vol. 143, October 1989, pp. 1234-1239.

Millstein, Susan G. "The Potential of School-Linked Centers to Promote Adolescent Health and Development." Washington, DC: Carnegie Council on Adolescent Development, August 1988.

National Education Goals Panel. "The National Education Goals Report 1991. Building a Nation of Learners." Washington, DC: U.S. Government Printing Office, 1991.

Newton, Jerry. "Primary Health Care in a School Setting." *Journal of School Health*, vol. 49, no. 1, January 1979, p. 54.

Novello, Antonia C., Christopher DeGraw, and Dushanka Kleinman. "Healthy Children Ready to Learn: An Essential Collaboration Between Health and Education." *Public Health Reports*, vol. 107, no. 1, January-February 1992, pp. 3-14.

Oldfield, D., and R. Petoss. "Increasing Student 'On Task' Behaviors Through Relaxation Strategies." *Elementary School Guidance and Counseling*, vol. 20, no. 3, 1986, pp. 180-186.

Olds, Robert Scott, and James M. Eddy. "Negative Health Messages in Schools." *Journal of School Health*, vol. 56, no. 8, October 1986, pp. 334-336.

Pate, Russell R. "Health-Related Measures of Children's Physical Fitness." *Journal of School Health*, vol. 61, no. 5, May 1991, pp. 231-233.

Pentz, Mary Ann, Bonnie R. Brannon, Ventura L. Charlin, Edith J. Barrett, David P. MacKinnon, and Brian R. Flay. "The Power of Policy: The Relationship of Smoking Policy to Adolescent Smoking." *American Journal of Public Health*, vol. 79, no. 7, July 1989, pp. 857-862.

Pentz, M. A., and P. Tolan. "Social Skills Training with Adolescents: A Critical Review of Time Trends, Dimensions, and Outcomes, 1972-1982." Manuscript submitted for publication as reported in Gilbert J. Botvin and Thomas A. Wills, "Personal and Social Skills Training: Cognitive-Behavioral Approaches to Substance Abuse Prevention." In *Prevention Research: Determining Deterring Drug Abuse Among Children and Adolescents*, edited by C. S. Bell and R. Batties. Washington, DC: Department of Health and Human Services, National Institute of Drug Abuse, 1985, pp. 8-49.

Polit, Denise, Janet Kahn, and David Stephens. "Final Impacts from Project Redirection." New York: Manpower Demonstration Research Corporation, April 1985.

Pollitt, E., M. Gersovitz, and M. Garguilo. "Education Benefits of the United States School Feeding Programs: A Critical Review of the Literature." *American Journal of Public Health*, vol. 68, no. 5, 1978, pp. 477-481.

Rosenberg, Morris. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press, 1965.

Rotter, Julian B. "Generalized Expectancies for Internal Versus External Control of Reinforcement." *Psychological Monographs*, vol. 80, no. 1, 1966, pp. 1-28.

Rowe, Daryl E. "Healthful School Living: Environmental Health in the School." *Journal of School Health*, vol. 57, no. 10, December 1987, pp. 426-431.

Rumberger, R. "High School Dropouts: A Review of Issues and Evidence." *Review of Educational Research*, vol. 57, no. 2, summer 1987, p. 101.

Schaps, Eric, Joel M. Moskowitz, Janet H. Malvin, and Gary A. Schaeffer. "Evaluation of Seven School-Based Prevention Programs: A Final Report on the Napa Project." *The International Journal of the Addictions*, vol. 21, nos. 9 & 10, 1986, pp. 1081-1112.

Schultz, Edward W., Raymond M. Glass, and J. Dennis Kamholtz. "School Climate: Psychological Health and Well-Being in School." *Journal of School Health*, vol. 57, no. 10, December 1987, pp. 432-436.

Silver, H.K., et al. "The School Nurse Practitioner: Providing Improved Health Care to Children." *Pediatrics*, vol. 58, no. 4, October 1976, pp. 580-584.

Simkin, Linda. "Evaluation: A New York Assessment of New York City High School-Based Clinics." Report to the New York State Department of Health. Albany, NY: Welfare Research, Inc., 1987.

Simons-Morton, Bruce G., Guy S. Parcel, Tom Baranowski, Ronald Forthofer, and Nancy M. O'Hara. "Promoting Physical Activity and a Healthful Diet Among Children: Results of a School-Based Intervention Study." *American Journal of Public Health*, vol. 81, no. 8, August 1991, pp. 986-991.

Stone, Elaine J. "Access: Keystones for School Health Promotion." *Journal of School Health*, vol. 60, no. 7, September 1990, pp. 298-300.

St. Pierre, Robert, M.K. Fox, M. Puma, F. Glantz, and M. Moss. "Child Nutrition Program Operations Study: Second Year Report." Cambridge, MA: Abt. Associates, January 1992.

Tesiny, Edward P., Monroe M. Lefkowitz, and Neal H. Gordon. "Childhood Depression, Locus of Control, and School Achievement." *Journal of Educational Psychology*, vol. 72, no. 4, 1980, pp. 506-510.

Thomas, Patricia A., and Margaret S. Texidor. "The School Counselor and Holistic Health." *Journal of School Health*, vol. 57, no. 10, December 1987, pp. 461-464.

University of Chicago, National Opinion Research Center. *High School and Beyond: Sophomore Questionnaire*. Baseline Instrument. Washington, DC: National Center for Education Statistics, 1980.

U.S. Department of Education. National Center for Education Statistics. *A Draft Proposal for a Longitudinal Studies Program of Young Children*. Washington, DC: National Center for Education Statistics, May 1992.

U.S. Department of Education, Office of Educational Research and Improvement. *The Effectiveness of Chapter 1 Services*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 1986.

U.S. Department of Health and Human Services. "Fifth Report to the President and Congress on the Status of Health Personnel in the United States." Washington, DC: U.S. Government Printing Office, March 1986.

U.S. Department of Health and Human Services, Public Health Service. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*. DHHS Publication No. (PHS) 91-50212. Washington, DC: U.S. Government Printing Office, 1990.

U.S. General Accounting Office. *School Dropouts: Survey of Local Programs*. HRD87-108. Washington, DC: U.S. General Accounting Office, July 1987.

Wang, M., M. Bear, J. Conklin, and R. Hoepfner. "Report 10: Compensatory Services and Educational Development in the School Year." Santa Monica, CA: System Development Corporation, 1981.

Waszak, Cynthia, and Shara Neidell. "School-Based and School-Linked Clinics: Update 1991." Washington, DC: The Center for Population Options, 1991.

Wells, L. Edward, and Gerald Marwell. *Self-Esteem*. Newbury Park, CA: Sage Library of Social Research, 1976.

Wigdor, Alexandra K., and Wendall R. Garner (eds.). *Ability Testing: Uses, Consequences, and Controversies. Part I: Report to the Committee*. Washington, DC: National Academy Press, 1982.

Wilson, N., and J. Rotter. "Anxiety Management Training and Study Skills Counseling for Students on Self-Esteem and Test Anxiety and Performance." *School Counselor*, vol. 34, no. 1, 1986, pp. 18-31.